Environment for Modular Image Reconstruction Algorithms and data analsysis (EMIRA)

MANUAL OF OPERATIONS

imcfromppcaem	Convert the ppcaem output coordinate file to
imodcoordinates2doc	Convert 3D coordinate file from IMOD to SPIDER-style
mask	Apply a round mask to an image or volume
mode	set some behaviors of EMIRA
model	Create model image (2D)
montage	montage many images into one large one
msa	Run multivariate statistical analysis
msaimc2doc	imc2doc Convert image coordinate file created by MSA to document file
msamap	Print 2D maps from PCA or coran coordinates.
msamaptoimage	create an image containing an MSA image coordinate map
msamaptoline	Projects the points from a 2D MSA map onto lines
msanlm	Apply nonlinear mapping to the image coordinates after msa.
msaselectbymask	select images according to mask of map
msavismap	Create visual representation of PCA map
multiply	Multiply two images
multiplyconjugate	Multiply two Fourier transforms
new2olddoc	Convert new format document file to old format document file,
oldimc2doc	Convert IMC image coordinate file from older coran to doc file
pba3	Peak file averaging for projection based 3D alignment
peak3	Find the maximum in a 3D file.
pickslice	pick a slice from a volume
pickstickturn	Program to box out the projection series of an
plothisto	plot a histogram from a document file to a postscript file
plotlines	lines Plot lines from a document file.
plotmlines	Plot lines from multiple document files.
plotscatter2d	Create a 2D scatter plot
plotscatter3d	Create a 3D scatter plot
ppcaem	PROBABILISTIC PRINCIPLE COMPONENT ANALYSIS WITH EXP. MAX.
predictresolution	predict resolution for conical or single axis 3D reconstr.
project	project a 3D volume onto 2D projection
qvol	vector quantization of volume (SITUS program)
rad2d	Calculate a 2D Redon transform
rad2d3dali	align a volume to a reference projections
rad2dinv	Invert 2D Radon transform
rad3d	Calculate a 3D Radon transform
rad3dcopycounter	copy counter of 3D Radon transfomr
rad3dinv	Invert 3D Radon transform
radalign2d	2D image alignment with Fourier-Radon transforms
radalignto3d	2D Radon transform alignment to 3D Radon transform
radextract	extract 2D Radon transform from 3D Radon transform
radfour	Calculate the 1D Fourier transform of a 2D or 3D Radon transform
radmake3d	Create an empty 3D Radon transform (or Fourier-Radon
radsigstat	Calculate histogram of sigmas in Fourier Radon transforms
radsum3d	Average 2D Radon transforms into 3D Radon transforms.
readdoc	read values from a document file
readdocclose	Close document file to free up memory

removeoutliers	outliers Remove outliers from document file columns
rot	rotate an image
rot3d	rotate a volume
scalevalues	Scale density values in an image
setangle	Put Euler angles in the header of an image.
shift	shift an image or volume
sqroot	square root of an image
square	square an image
subtract	subtract 2 images
table2doc	convert whitespace seprated table to SPIDER document file.
threshold	Theshold an image
variables	their use in EMIRA
vismap	Create a visual map of msa results
vs2doc	convert image coordinates from an XMIPP neural net to
window	Box out small image or volume from large image or volume
writedoc	write numbers into a document file
writedocclose	Close document file to make is useful for reading.
zeiss2spider	convert Zeiss/SCAI scanner tif to spider format

add add 2 images

file format: spider

USAGE: add

.Input file: image001
[Enter input file name]

.Second input file: image002
[Enter name of seconf image]

.Summed Output file: image003
[Enter name of output file containing the sum of both images.]

Programs: add.f (extracted from spider)

Author(s): M.Radermacher

appenddoc **append an ascii file to another**

file format: any ascii files

PURPOSE: Append a document file to the end of another one. This is mainly
used for selection files when additional particles need to be excluded.

USAGE: appenddoc

.Input file: excludedoc001
[Enter input file name.]

.Output file: select001
[Enter name of output file.]

REMARK: This program uses the generic "cat" command combined with '>>'

Programs: cat (shell command)

Author(s): M.Radermacher

ask **Ask for input values in a procedure**

ask filename
ask number
ask text

CAUTION: THIS MECHANISM IS NOT EXTENSIVELY TESTED YET, USE WITH CAUTION.

USAGE:

General: All ask commands will create a prompt at the input level one higher than the procedure. This means, if the procedure is called interactively, the prompt will go to the terminal. If the procedure is called from a batch file, the answers are read from the batch file in the lines following the procedure call. In a bacth file, the answers must be preceeded by "<". This is to avoid that they are not processed again as commands (which most probably will not exist). i.e. the "operation" prompt in emira treats these lines as comments.

IMPORTANT, THE ITEMS IN THE LINE MUST BE SEPARATED BY COMMAS !!!

askfilename, v.file, Please enter file name
[Use the command followed by a comma separated list. v.file is a user specified variable that will receive the file name. The next item is the prompt that will be echoed to the input. The file name question can be answered like the file name questions in single commands. For example it can be file[3] v.number, where v.number contains a file number, which is substituted for the [3]. If a file is specfied as file[3] and no number is provided, then "file[3]" will be the value v.file. This may be needed if the procedure needs to know where to substitute the file number. A mixture like file[3][4] v.num is currently not allowed. Either all numbers need to be substituted or none.
If a variable, for example v.inputfile contains the file name, this also can be used as answer, if applicable with the same number substitution rules.]

asknumber, v.varaible, Please enter value
[Use the command followed by a comma separated list. The first item will receive the number that was answered, the second item is the prompt. Variables may be used to answer the prompt.]

asktext, v.text, Please enter yes or no
[Use the command followed by a comma separated list. The user specified variable v.text will receive the answered text. The answer can also be contained inside a variable.]

average calculate averages of an image series

```
USAGE:        average

        .First input filename: image001
        [Enter the name of the first image that will be averaged.]

        .Add to existing file Y/N: N
        [Specify if you are adding images to an already existing average.]

        If the answer to the question above was "Y"
                .Number of images already averaged: 127
                [Enter the number of images that were used to calculate the
                existing average.]

                    .Average offset from before: 0.23
                    [Enter the average offset that was determined by the previous
                     averaging step.]

        .File numbers in document file? (D): D
        [Answer "D" if the file numbers should be read from a document file.]

        If "D" was answered:
                .Document file: Numdoc001
                [Enter the name of the document file that contains the file
                 numbers.]

                .Column containing file numbers: 3
                [Enter the column that contains the file numbers. (key=0).]

        Otherwise:
         .File numbers: 1-250
         [Enter the file numbers of the images that are averaged. Missing images
          allowed.]

         .Average all (def), split odd/even, (A/O): O
         [Enter if you want to average all images of split them into two sets
          and calculate two averages and veriances. Splitting the data set
          is intendend for resolution measurements.]

         If "A" was answered:
         .Average file name: average001
         [Enter name of the average image.]

         .Variance file name: vari001
         [Enter name of the variance image.]

         If "O" (split odd/even) was answered:
         .Odd vaverage file name: odaverage001
         [Enter file name for the average of the odd particles.]

         .Odd variance file name: odvari001
         [Enter file name for the variance image of the odd particles.]

         .Even average file name: evaverage001
         [Enter file name for the average of the even particles.]

         .Even variance file name: evvari001
         [Enter file name for the variance image of the even particles.]

         .Enter varaible of output odffset: v.off
         [Enter a variable that will receive the offset value of the average
          witbin a named tuple. The offset will be found in v.off.offset]

Programs: adds_m.f, in /extras/fstandalone/applications. Program
          estracted /adapted from SPIDER.

Author(s): M.Radermacher, others.
```

averagehistogram calculate histogram for all point in an average

```
image format: SPIDER

USAGE:       average histogram

        .First input filename: image001
        [Enter name of first image file used for the average (2D only).]

        .File numbers in document file? (D): D
        [Enter 'D' if the file numbers are provided in a document file.]

        if D was answered:
                .Document file: filenumberdoc001
                [Enter name of document file that contains the list
                 of file numbers.]

                .Column containing file numbers: 3
                [Enter which column in the document file contains the
                 file numbers. (key and the next value do not count).]

        If anything else than D was answered:
                .File numbers: 1-10,30-300
                [Enter the file numbers.]

        .Histogram file name: avghisto001
        [Enter the name of the output file. This is a 3D file, where the
         histogram of each point is along the X-axis. The x-dimension depends
         on the number of bins selected. The y,z dimensions are the original
         x,y dimensions.]

         .Lower, upper bounds: 0.1,1.0
         [Enter lower and upper value bounds for the histogram.]

         .Number of bins: 20
         [Enter number of bins used in the histogram.]


Program: histoser.f (part of core)
Author:  M. Radermacher
```

backgrounds determine averages of subareas

PURPOSE: Determine the average background of an image that is uneven.
The program picks 4 square areas (at this time) at a specified
distance in x, -x, y, -y from a user defined center and determines
the median of the average values.

USAGE: backgrounds

.Input file: image001
[Enter the name of the large input file.]

.Center x,y, distance from center: 257,257,50
[Specify image center, and how far the four subimages should
be away from this center (cento to center measurements).]

.Box dimension: 50,50
[Enter the dimensions of the boxed out subimages.]

.Enter variable to receive output: v.avg
[Enter the varaible name for a named tuple to recieve the median
average and sigma. Default name v.back.]

.First output file or *: litwin001
[Optionally provide an output file name and the four windows will be
written out. Numbering will be from 1-4.]

Programs: subimstats.f,

Author(s): M.Radermacher

backproject calculate a simple backprojection

image format: SPIDER

USAGE: backproject

```
          .3D volume: volume001
          [Enter the name of the 3D output volume (or input volume if you are
          adding to an existing one).]

          .Is this a new 3D: Y
          [Enter if this is a new 3D to be created (Y), or an existing 3D that
           you want to add to (N or any other letter).]

          If it is a new 3D:
             .x,y,z-dimension: 2048,2048,800
             [Enter the dimension of the volume to create.]

          .ix,iy,iz center coordinate: 0,0,0
          [Enter the coordinates for the center if you are calculating
           a subtomogram.
           The x,y coordinates are relative to the 0 degree
           projection, using the conventions of "pixel" in SPIDER WEB. I.e.
           x and y start at array coordinates (1,1).
           z is the location of the center of the volume. If 0,0,0 is answered
           the origin of the 3D is placed at te default center:
           (x-dim/1+1, y-dim/2+1, z-dim/2+1.)]. If z was determined in imod,
           then the answer here should be: Ds/2-Do/2+Zm, where Ds is the
           z-dimensions of the small volume, Do is the z-dimension of the
           volume from which the coordinate was picked, Zm is the picked
           z-coordinate.

           .Enlargement: 1.0
           [Enlargement allows for different magnifications in 3D and 2D. It
            also may be smaller than 1 to create a size reduced reconstrution.
            Note however, that demagnification may create a noisier
            reconstruction because of possible subsampling.]

           .Name of first projection: proj001
           [Enter the name of the first projection in the series. The
           projections must have the projection angles in the header in header
           position 1. They may contain all 3 Euler angles.]

           .First, last projection number: 1,72
           [Enter the first and last projection number.]
```

Notes: The fortran programs are extensions of the original backprojection written 1978.(see M. Radermacher, W. Hoppe, 3-D Reconstruction from Conically tilted Projections. Proc. 9th Int. Congr. El. Micr. Toronto 1978, Vol I 218-219 and M. Radermacher, thesis, 1980). For each projection the algorithm loops through the volume and adds the interpolated projection value to the 3D volume. Modifications include the extension to 3 Euler angles (instead of 2 angles in the original), and the allocation of the volume in-core (instead of one slice at a time).

For a large volume make sure that the memory of your computer is sufficiently large to avoid usage of swap space. Rule of thumb, the available memory should be at least 1.3 times the size of the volume. If memory is insufficient the volume should be reconstructed in chunks, which can be done by using the subvolume options.

Programs: em_backproject.py, bapi3absubv.f, mirbp3absub.f

Author(s): M. Radermacher

bconvert Convert image formats

image format: many

USAGE: bconvert

  .Input file: image001
  [Enter the name of the input file. Note, that if you do not
  provide an extension the data code will be attached.]

  .Output file: converted001.mrc
  [Enter the name of the output file. It is a good idea if the
  extension is provided and matches the file type. In any case
  the output file will get this name independent of the format
  of the content.]

 The program will write:
  Current input file extension: st1
  [Assuming that .st1 was your data code for this session,
  and thus the extension of the input file.]

  .Enter extension that specifies the file format or or ?: spi
  [Enter an extension that describes the input file format. You may hit
  return if the extension matches the format. If you enter "?" you
  get a list of all possible answers/formats.]

 The program will write:
  Current output file extension: mrc
  [Lists the extension you used for the output file.]

  .Enter extension that specifies the file format or or ?: mrc
  [Enter an extension that describes the input file format. You may hit
  return if the extension matches the format. If you enter "?" you
  get a list of all possible answers/formats.
  NOTE: THE FORMAT SPECIFICATION WILL NOT CHANGE THE OUTPUT FILE NAME
  EXTENSION. THIS WILL BE THE ONE YOU USED IN THE ANSWER TO "OUTPUT FILE"]

Programs: em_bconvert.py, bconvert.c and suport programs from BSOFT.

Authors: M. Radermacher (python), B. Heyman et al. bconvert and support programs.

Remarks: This command uses the unchanged version of bconvert from BSOFT.
  The decision, that the file format specification does not change the specified filenames
  was made, because the SPIDER image extension can be any 3 letter combination and does
  not specify the file format, but the data set. Therfore there can be a mismath between
  format and extension in either the input ot the output file. (If you want to confuse
  yourself you could use tif as the spider data extension. This does not make the content
  tif files.)

File formats supported (if not please contact B. Heyman):

FORMAT:	EXTENSIONS:
ascii	asc txt
Biorad	pic
BRIX	brx brix
Broohaven STEM	dat
CCP4	ccp map
Digital Instruments atomic force	di
Digital Micrograph (GATAN):	dm
Alwyn Jone's map format	omap dsn6 dn6
EM system format	em
Peter Goodford's GRID files	pot
Basel GRD map file format	grd
HKL file typically produced with MTZDUMP from X-ray crystallographic structure factor files in the CCP4 package.	hkl
IMAGIC	img hed
Digitabis image plate reader	ip
JPEG:	jpg jpeg

text kernel file:	krn
3D file "What If" molecular dynamics package.	mif mff
MRC format(including IMOD ext.)	mrc stk st ali rec
Bsoft default format	pif sf
PIC package, NIH	bp bq
Portable Network graphic	png
portable bitmap, grey map, pixel map	pbm pgm ppm
Princeton Instruments CCD image file format.	spe
spider	spi
suprim new format	spm sup f
Tiff: 8bit only	tif
XPLOR map or structure factor file formats.	xpl cns rfl

bin Bin down image

image format: SPIDER

SEE ALSO: bin density (Bin down image and convert to optical desities)

USAGE: bin

```
.INPUT FILE: image4086
[Enter the name of the image containing transparencies]

.OUTPUT FILE: red4086
[Enter name of output file]

.Size reduction factor for binning (def. 1): 3
[Enter a binning factor if you want to reduce the image size and
 increase the pixel size. Note: in some cases binning can cause
 aliasing. Normally though the true resolution of the image is
 lower that the Nyquist limit of the scan.]

.Scale factor of density (def 1.): 1.0
[Enter an optional scaling factor for the density values.]
```

Notes: The program bins down an image by an integer factor. The new dimension is calculated byinteger division. If there is a remainder, the corresponding pixels at the end of the rows and columns in the original image are ignored.

Programs: em_bin.py, mr_ipavg2.f

Author(s): M. Radermacher

bin3D Bin down volume

image format: SPIDER

SEE ALSO: bin density (Bin down image and convert to optical desities)
bin (bin down 2d image)

USAGE: bin 3D

.3D Input file: volume001
[Enter the name of the input volume]

.3D output file: redvolume001
[Enter name for the binned down output file]

.Size reduction rx,ry,rz (def. 1): 3,3,3
[Enter a binning factor if you want to reduce the image size and increase the pixel size. Note: in extreme cases binning can cause aliasing. Normally though the true resolution of the volume is lower than 2 sampling steps. If there is a concern, first low-pass filter the volume, then bin it down.]

.Scale factor of density (def 1.): 1.0
[Enter an optional scaling factor for the 3D density values.]

Notes: The program bins down a volume by an integer factor. The new dimensions are calculated by integer division. If there is a remainder, the corresponding pixels at the end of the rows and columns in the original image are ignored.

Programs: em_bin3d.py, mr_ipavg3.f
Program group: Core

Author(s): M. Radermacher

bin density Bin down image and convert to optical densities

SEE ALSO: conv nikon (convert NIKON Supercoolscan tif to
                spider format)
           conv scai (Convert tiff file from Zeiss SCAI scanner)

USAGE: BIN DENSITY or bindensity

          .INPUT FILE: image4086
          [Enter the name of the image containing transparencies]

          .OUTPUT FILE: raw4086
          [Enter name of output file]

          .Lower / upper threshold of histogram during scan (def. 0,1): 0,1
          [If thresholds were applied to the image histogram during scanning, enter the
           values here. 0,1 is the full histogram range.]

          .White value from reference scan (def 65535.): 4000
          [Enter the brightness measurement from a reference scan of a blank film area.
           The maximum possible values are: 256 for 8-bit scanners, 1024 for 10-bit
           scanners, 4096 for 12 bit scanners and 65636 for 16 bit scanners.]

          .Size reduction factor for binning (def. 1): 3
          [Enter a binning factor if you want to reduce the image size and increase the
           pixel size. Note: in some cases binning can cause aliasing. Normally though
           the true resolution of the image is lower that the Nyquist limit of the scan.]

          .Scale factor of density (def 1.): 1.0
          [Enter an optional scaling factor for the density values. Normally this should
          be 1.0, but if you want to invert contrast you may enter -1.0 for example. Scaling
          is applied after the OD conversion.]

Notes: The program calculates the LOG10(I0/I), I0 being the value supplied to the question
          of the white reference. (At this time there is no provision for dealing with negative
          output values,which may occur if part of the empty glass was scanned or if the white
          value entered is too small. If emty glass was scanned, then windowing the input image
          to a smaller size should eliminate the negative values.
          This program was originally developed to convert data from either the scai scanner (12 bit)
          or Nikon scanner (16 bit). First, the scanner output is converted to a spider format
          image without any change in values. This program then calculates optical densities and
          can be used to increase the pixel size by binning.

Programs: em_bindensity.py, mr_ipavgdens2.f

Author(s): M. Radermacher

calcproshifts calculates projection shift based on 3d volume shift

```
file formats: SPIDER document file

USAGE:       calc proshifts

        .Input document file: proangles001
        [Enter name of document file that contains the projection angles.]

        .(O)ld or (N)new docfile format (default N): N
        [Enter O if the input document file has the old formatted SPIDER
         format. The new format is space separated columns, while the old
         format is formatted and it can happen that there is no white-
         space between the columns. In this case, trying to read with option
         N will give a formatted access error in the fortran code.]

        .Columns containing phi, theta, psi: 3,4,5
        [Enter the column numbers that contain the angles. Use document
         file convention (ignore the first two columns containing key and
         number of items).]

        .Output document file: proshiftdoc001
        [Enter name for document file that will receive the shifts for each
         projection.]

        .X-shift,Y-shift,Z-shift: 10.2,12.1,-5.0
        [Enter 3D translation.[

Programs: calcproshifts.f
Author: Michael Radermacher, 2014
```

calslope **calculate the tunning slope of a curve, averaged over a** specified interval.

file formats: SPIDER document file

PURPOSE: This program was written to evaluate the slope of a spectrophotometer curve. It runs an interval along the curve and determines the slope within the interval as a linear least squares fit. The interval is moved along the curve and the output contains a function that shows the slope in the center of the interval at everyt position.

USAGE: calc slope

.Input document file: functiondoc001
[Enter the document file that contains the function. (see also command tabe2doc for transferring a space separated table into a document file)

.Output doc file: slopedoc001
[This file will contain the following columns (after the key and counter): X,Y of input function, m=slope*factor, m*1, n intersect, error of slope*fact, error of n, standard deviation over the interval.]

.x-, y-column, interval: 2,5, 25
[Enter the column for the x and y of the function, and specify the interval length.Columns are counted for spider document files. I.e. the first two columns (key and itemnumber) are ignored.]

.highest key to use: 121
[Enter the last key in the input document file that should be used for the calculation. (This question was introduced because the curves stored in the spectrophotometer files may be padded with 0, which causes the slope to be undefined.)]

.factor for slope: 1.0
[Enter a factor for the slope. This can in certain cases give you the value in the units you want.]

Programs: em_calcslope.py, calcslope.f
Author: M. Radermacher
Author:

calczinplane Calculate a z-value in a plane based on the x,y coordinates
and the paramters of the plane equation.

file format: SPIDER document file

SEE ALSO: fitplane

PURPOSE: This program was developed to calculate the defocus in any point in a tilt image. The plane equation is calculated with fitplane from the CTF fitting procedure, and then used in this program to determine the defocus in the center point of each single boxed out image. The correction then is applied to each individual particle image.

USAGE: calczinplane (or calc z in plane)

.Input doc file with x,y coords: coordinates001
[Enter the name of the input document files that contains the x,y coordinates.]

.Enter columns of image-number, x,y: 0,1,2
[Enter the columns where to find the image number (or coordinate number) and the x and y coordinates.]

.Output doc file: zcoordsdoc001
[Enter output document file that will receive the resulting z-coordinates (s.also below).]

.Enter columns for image-number,z,scaled z: 1,2,3
[Enter which column should contain the image/coordinate number, the calculated z-coordinates and the z-coordinated scaled with the pixel size. Note: Most of the time this program has been used to calculate the defocus in a tilted image. Since in this case the z-coordinate is already in Angstroems, the scaled z-value is irrelevant.]

.Enter plane prameters a,b,c: -1.2,3.0,-2100.
[Enter the paramaters of the plane equation. These may have been determined with fitplane.]

.Enter scaling factor: 4.02
[Enter a factor for scaling the z-value. Somtimes this is useful, other times not. defaults is 1.0]

Programs: calcplanez.f, and support routines.
Author: M.Radermacher

chim2euler Convert UCSF Chimera Matrix to EMIRA document file

file formats: see below

USAGE: chim2euler

File containig matrix from chimera: chimeramatrix.txt
[Enter the name of the file created by "matrixget" with Chimera]

Output document file: doc001
[Enter the name of the document file to be created. This file
will have the format: key, 6, phi, theta, psi, xs, ys,zs
The key corresponds to the volume numbering that chimera assigns +1.
The +1 is needed for compatibility with SPIDER since the spider
document file cannot have a key smaller than 1.]

Chimera number of reference volume, Volume dimensions x,y,z: 0,160
[Enter the number that Chimera assigned to the volume that was
used as a reference to align the other volumes. Enter the dimen-
sions of the volumes. If only one or two dimensions are provided,
the other dimensions are set to the x-value. The dimensions are used
to calculate the volume origin and then the shifts that need to
be applied.]

Programs: em_chim2euler.py, chim2euler (binary), chim2euler.f (in fstandalone/
applications)

Author(s): M. Radermacher, 2012

classi diday Classification with moving centers.

file formats: input: special, output: SPIDER document file

PURPOSE: Carry out a classification after correspondence analysis or
Principal component analysis, using Diday's method of moving
centers. (was cl cla in SPIDER 5.0)

USAGE: classi diday

.Prefix for input coordinate file: first
[Enter the file prefix that was used in msa (coran,PCA). This
created the files prefix_IMC,prefix_EIG,prefix_PIX. Classi diday
uses prefic_IMC.]

.Output Cluster file: cluster001
[Enter the name of a (binary) cluster file, that classi diday
creates.]

.Factor numbers to use: 1-3,6
[Enter the factors you want to use in the classification.]

.Number of iterations, centers, partitions: 5,3,5
[Enter the number of centers you want to use for each partition.
Default is 3,3,3]

.Class cutoff %: 0
[Specify the minimum class membership as a percentage of the total
number of data points. Classes with fewer members than this cutoff
are ignored.]

.Output dendrogram docfile: dendrodoc001
[Enter a name for the document file that will recieve the images (as
key), followed by the linkage height.]

.Output cluster doc: clusteroc001
[This will contain a readable version of the cluster file.]

Remarks: The Rsults file will contain the following information:
A) Aggregation history:
For each of the NP partitions, the NS seeds and sizes of
clusters are listed. The non-empty clusters in the crossed
partion are given by size and cummulative percentage. A
total of NS**NP clusters are possible, but in practice, only
10% of these are non-empty.

B) Description of hierarchy nodes in HAC:
The nodes are numbered starting from the number of the highest
cluster. For each node, the Senior, Junior, size, weight
and hierarchy index are given. The hierarchy index is printed
out as a histogram.

C) Description of the hierarchy classes:
For each node the constituent classes are listed.

D) HAC Dendrogram:
The class relationships are represented in the form of a
dendrogram (tree structure). The lengths of the branches (in
horizontal print direction) are proportional to the hierarchy
indices.

E) List of class members:
The members are listed for each of the basic non-empty classes.

F) List of class center coordinates:
For each class, the NFAC coordinates of its center is listed.
These are contained in the cluster file.

G) Re-classification lookup table:
Each cutoff point in the dendrogram, from right to left, defines
a new classification scheme with the number of classes increasing
by 1 each time. The table gives the new class memberships for
any cutoff point selected.

2. The algorithm underlying this set of programs is described by
   Lebart et al., MULTIVARIATE DESCRIPTIVE STATISTICAL ANALYSIS,
   J. Wiley and Sons, New York 1984, Chapter V, p.109.

Programs: SCLASSI, SCLASSY, SEMIS, NOYAU, CHAVA, DEUCL, RGRI,
RCLU, DENDRO, ARBRE, PARST, STABK, SHELK, COUPE, CLAST,
RETIR, SEN3A, CLUSTERLST

Authors: M. Radermacher, J.P. Bretaudiere

combine euler Combines 2 Euler rotations into a single one

file formats: output: SPIDER document file (no input file)

USAGE: combine euler

```
        .Enter phi1, theta1 and psi1: 10.0, 90., -30
        [Enter the three angles of the first Euler rotation,
         phi rotation around z, theta rotation around the new y,
         psi rotation around the new z. Theta rotation is negative,
         consistent with definitions of Euler angles also in SPIDER.]

        .Enter phi2, theta2 and psi2: 20.,-30.,20.
        [Enter the second set of Euler angles.]

        .Enter varaible to receive output: v.ang
        [Enter the variable that will contain the output values. This
         will have the form of a named tuple. Access the single values
         with v.ang.phi, v.ang.theta and v.ang.psi. You can also print
         the tuple to see what is in there. (print v.ang).]

        .Create output document file Y/N: Y
        [ You can create an output document file that will contain the
         euler matrix of the combined rotation. On line contains 9
         numbers, each three form one column of the rotation matrix.
         column 1 : value 1-3 column 2: values 4-6, column 3: values 7-9]

         If 'Y' was answered then
            .Enter document file name: matrixdoc001
            [Enter the name of the document file that will contain the
             resulting matrix.]

            .Enter key: 10
            [Enter the key under which the matrix should be stored in the
             document file.]
```

Remarks: This is a rewrite of the formare SPIDER vo ceul command for external input.

Combines two sets of three Euler angles into a single set of three Euler angles. The definition of the Euler rotations is as follows:

r' = D(psi) D(theta) D(phi) r

with:

$$D(psi)=\begin{vmatrix} \cos(psi) & \sin(psi) & 0 \\ -\sin(psi) & \cos(psi) & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$D(theta)=\begin{vmatrix} \cos(theta) & 0 & -\sin(theta) \\ 0 & 1 & 0 \\ \sin(theta) & 0 & \cos(theta) \end{vmatrix}$$

$$D(psi)=\begin{vmatrix} \cos(psi) & \sin(psi) & 0 \\ -\sin(psi) & \cos(psi) & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

These definitions are consistent with the angle definitions used in "backproject" and the Radon 3D reconstruction programs.

Programs: em_combineeuler.py, combeuler.f combeul.f mrmamule.f

Author: M. Radermacher 1990 and 2013.

convertmarkers **convert marker selected in IMOD for tomography** alignment to a series of SPIDER format document files.

USAGE: convert markers

.Input IMOD marker text file: markers001.txt
[Enter the file created from the imod coordinates using the IMOD command "model2point -float -contour .fid markers001.txt" Please note that the input text file must contain the marker coordinates as floating point numbers. It is a text file with columns, where each line contains:
|marker number| |X-value| |Y-value| |projection number|
separated by white space.]

.First output doc file: mdoca001
[Enter the name of the first output SPIDER format document file. The program will create a series of document files, one for each projection.]

.First tilt angle, increment: 60.,3.0
[Enter the tilt angle of the first projection and the angular increment.]

REMARK: The program will create a series of document files containing the markers for each projection. Markers that are not present in all projections will be discarded. The markers are renumbered in sequence, and the original marker number will be listed as the last value in each line of the document file.

PROGRAMS: convertmarkers.f
AUTHOR: M. Radermacher

copy Copy a file (any format)

file format: any

USAGE: copy

.Input file: image001
[Enter input file name]

.Output file: subdir/immasked001
[Enter name of output file. If the directory does not exist, it is created].

REMARK: This program uses the generic "cp" command

Programs: cp (shell command)

Author(s): M.Radermacher

copyspider copy spider files (optional stacks)

file format: SPIDER

USAGE: copy spider

```
        .Input file: image001
        [Enter input image name. Input file can also be a slice of a stack,
         identify with "@" and number.]

        .Output file: outputfile001@001
        [Enter output image name. Directory will be created if needed.
         The answer may either be a simple file name or, if it contains
         the symbol @ a slice in a stack.]
```

remark: This operation does not copy whole stacks in one command. Use simple "copy" instead.

Programs: copyspider.f

Author(s): M.Radermacher, others.

createdir create a new directory

file format: N.A.

USAGE: create dir

```
        .Directory name:  ../newdir
        [ Enter name of directory]
```

REMARK: This program uses the generic "os.mkdir" command. If the directory exists, nothing happens.

Programs: os.mkdir

Author(s): M.Radermacher

crossco Cross-correlate two images, not normalized.

file format: SPIDER

USAGE: crossco

.Reference file: refer001
[Enter input image or volume.]

.Input file: image001
[Enter name of input file that is being compared to the reference.]

.Output file: ccf001
[Enter the name of the output file to contain the cross-correlation function.]

REMARK: This crosscorrelation is not normalized. For use in translational alignment a normalization is not needed.

Programs: crossco_m.f, and many subroutines in
/extras/fstandalone/applications. Program
estracted from SPIDER.

Author(s): M.Radermacher, others.

crossconorm Cross-correlate two images, normalized.

file format: SPIDER

USAGE: crossco norm

```
        .Reference file: refer001
        [ Enter input image or volume]

        .Input file: image001
        [Enter name of input file that is being compared to the reference.]

        .Output file: ccf001
        [Enter the name of the output file to contain the normalized
        cross-correlation function.]
```

Programs: crossco_m.f, and many subroutines in
/extras/fstandalone/applications. Program
estracted from SPIDER.

Author(s): M.Radermacher, others.

delete delete a file (any format)

file format: any

USAGE: delete

```
        .Input file: image001
        [ Enter input file name. For safety wild cards like * and ? are
          not allowed and will result in an error message.]
```

REMARK: This program uses the generic "/bin/rm" command

Programs: /bin/rm (shell command)

Author(s): M.Radermacher

display **display a 2D image**

file format: many (all that bconvert can handle)

USAGE: display

.Input file: image001
[Enter name of input file.]

.Enter extension that specifies file format or : spi
[Enter an extension that specifies the file format. e.g.
spi for SPIDER format. If the extension already specifies
it, you may hit return. For valid file format extension see
the manual for bconvert (which is used within "display").]

Remarks: The program uses bconvert to convert the image to .jpg and then
displays it using the gpl licensed program geequie. geequie allows
images to be send to a running instant of the program without
opening another instance (except if the timing is too fast, as may
occur when display is used in loops).]

Programs: adm_showpict.py, bconvert, geequie

Wrapper implemented by M.Radermacher

divide **Divide one image (volume) by another**

file format: SPIDER

USAGE: divide

.Input file (divident): zaehler001
[Enter the input file that is the divident in the division.]

.Input file (divisor): nenner001
[Enter the input file that contains the divisor.]

.Output file: fraction001
[Enter the name for the output file.]

Remark: If a 0 is encountered in the divisor, this point in the output image is set to 0.

Programs: em_divide.py, add.f (from spider)
Author(s): M. Radermacher and others (for parts of the fortran code)

doccalc Carry out calculations on document file columns.

file format: input: SPIDER document file

USAGE: doccalc

```
        .Input document file: doc001
        [Enter document file name.]

        .New (default) or old docfile format (N/O): N
        [Specify  if the document file follows the current (new) space
         separated column format, of the old strict format. If its the
         old format, continuation lines will be ignored though.]

        .Name for outpt variable: v.values
        [Enter the name for an output variable that will receive the
         results of the calculations. This is a python named tuple.
         Specific values returned, see below, depends on option.]


        the following arithmatic is possible:
                1. 1D statistics (min,max,average,sigma)
                2. 2D vector lenght statistics
                3. Average length & direction of polar vector
                4. Find threshold location descending

        .Specify type of calculation (1,2,3 or 4): 1

        if "1" was specified, calculate statistics of a single column:
            .Column number: 3
            [Enter column number in document file.]

        v.values will contain (example):
        docarith(min=-0.05496, max=0.99998, avg=0.260748, sig=0.395849)


        if "2" was specified, calculate statistics of a 2d vector
             with x in first column and y in second column.
            .Column numbers: 3,4
            [Enter column numbers in document file for the components of
             the 2D vectors.]

        v.values will contain (example):
        docarith(min=0.162052, max=1.00006, avg=0.493869, sig=0.276385)
          where min is minimum length, max is maximum length,
          avg is average length, sig is the standard deviation.


        if "3" was specified, calculate statistics of a 2d vector
             in polar coordinates.
            .Column numbers: 3,4
            [Enter column numbers in document file for radius and angles
             of the 2D vectors.]

        v.values will contain (example):
        docarith(x=0.260747, y=0.000331465, rad=0.260748, ang=0.072835)
          where x and y are the coordinates of the average vector, rad
          is the average length and is the average angle (Note: the average
          angle only is meaningfull, if the angular range in the document
          file makes sense. It is simply the straight average of the numbers
          the program finds. No adjustment for any redundancy.]


        if "4" was specified, calculate location of a threshold value
             determined in descending order. Purpose is to find the radius
             in a Fourier Ring Correlation file, where the curve crosses
             a certain threshold.
            .Check (e.g. FRC) column, coord. column: 3,1
            [Enter column numbers of the Fourier Ring correlation and
             specify the column number for the x-coordinate. The return
             will be the interpolated x-coordinate for the location  of the
             threshold value (s. below)]
```

```
            .threshold value: 0.3
            [Specify the threshold value, for which the interpolated
             corrdinate is determined.]

        v.values will contain (example):
        docarith(key=23.0, rad=0.139642, frc1=0.3612, frc2=0.18262)
        where key is the key after crossing the threshold, rad is the
        interpolated Fourier radius where the threshold is crossed (assuming
        that in the example column 1 contains the radius in absolute Fourier
        units), frc1 is the value above the threshold, frc2 the first value
        below the threshold. This example uses threshold 0.3. ]


PROGRAMS: docarith.f

Author(s): M. Radermacher
```

exposurelog2doc Convert the exposure log file from a tecnai microscope
to a spider-style document file.

file format: see below

USAGE: exposurelog2doc

.File containing log file from microscope: miclog_1819_to_3410.txt
[Enter the file name that contains the microscope log. Since the program is primitive, the microscope settings must be such, that the log file from the microscope must contain the following items in the exact order and no other items:
magnification, spotsize, defocus, HV, emission, plate#, tiltangle
If the order is different you may be able to change it, maintaining tab separated columns. Recommended is to set up the logging on the microscope such that this format is maintained. An example of the beginning of a microscope log file:

Exposure Log 2/1/2005 6:58:45 PM

Time	Mode	Magn/CL	Spot	Defocus	HT	Emission	Plate No.
2/16/2005 11:00:34 AM	Image	52000	5	-2.184	100	2	A01819
2/16/2005 11:01:25 AM	Image	52000	5	-2.184	100	2	A01820
2/16/2005 11:03:42 AM	Image	52000	5	-1.248	100	2	A01821
2/16/2005 11:07:13 AM	Image	52000	5	-2.184	100	2	A01822
2/16/2005 11:09:19 AM	Image	52000	5	-1.326	100	2	A01823
2/16/2005 11:12:08 AM	Image	52000	5	-2.262	100	2	A01824
2/16/2005 11:14:53 AM	Image	52000	5	-1.248	100	2	A01825

]

.Output document file: exposuredoc001

Programs: em_exposurelog2doc.py, in direcory /extras/fstandalone/converters:
emira_exposurelog2doc.f

Author(s): M. Radermacher

filestats Get file statistics (Min,MAx,AV,SIG)

file format: SPIDER

USAGE: filestats

.Input file: image001

.Enter variable to receive output: v.stats
[The output variable will receive a named tuple. To look
at it use the command print, to use the results acces the
return values as: v.stats.min, v.stats.max, v.stats.avg and
v.stats.sig. These can be used like normal variables. If no
answer was given for the output variable, then the name
v.filestats is used by default.]

Programs: filestatistics_m.f, in /extras/fstandalone/applications. Program
estracted from SPIDER.

Author(s): M.Radermacher, others.

filestatsmask Get file statistics in mask (Min,MAx,AV,SIG)

file format: SPIDER

USAGE: filestats mask

.Input file: image001
[enter input file name]

.Mask file: mask001
[Enter name of a mask file. The statistics is determined in areas where the mask image has a value greater than 0.5]

.Enter variable to receive output: v.stats
[The output variable will receive a named tuple. To look at it use the command print, to use the results acces the return values as: v.stats.min, v.stats.max, v.stats.avg and v.stats.sig. These can be used like normal vaiables. If no answer was given for the output variable, then the name v.filestats is used by default.]

Programs: filestatistics_m.f, normm.f in /extras/fstandalone/applications.
Programs estracted from SPIDER.

Author(s): M.Radermacher, others.

findfourdimensions **Find next higher/lower dimension f. Fourier transf.**

file format: N.A.

PURPOSE: This program determines the next lower (or higher) value for image dimensions usable in Fourier transforms. The SPIDER Fourier transform, which is also used in many programs in EMIRA required the dimensions to be a product of prime numbers smaller than 7 (officially 23, but in reality this does not work). If an image has dimensions that are not usable in a Fourier transform this program will return the closest values of dimensions that will work.Images then can either be boxed out to the lower dimensions or padded to the higher dimensions.

USAGE: find four dimensions

Enter current x,y and z dimensions: 71,51,73
[Enter the current dimensions. If less then 3 values are entered, the others are set to 0.]

Enter variable to receive output (def: v.fourdim): v.newdims
[Enter a variable that will receive the named tuple containing the return values. If no answer is given the variable v.fourdim will be used.]

In the above example the program will return in v.newdims:
findfnum(xlow=70.0, xhigh=72.0, ylow=50.0, yhigh=54.0, zlow=72.0, zhigh=75.0)
which can be used as, for example: v.newxdim=v.newdims.xlow etc.

Programs: em_findfourdimensions.py, callfindfnum.f, finhigher.f, findlower.f
primedecomp23.f

Author: M. Radermacher

fitplane Fit a plane to a set of x,y,z coordinates

file format: SPIDER document file.

SEE ALSO: calczinplane

PURPOSE: The program was developed in conjunction with CTF correction in tilt images. Given the defocus value at several points in the tilt image, a plane can be fitted that describes the defocus for every point in the image. This in turn can be used to correct the CTF of individual boxed out particles, based on their location in the image. (Since the program is not specific for defocus only, there may be other applications for this program.)

USAGE: fit plane

```
.Input doc file with x,y,z coords: coordinates001
[Enter the name of the input document files that contains the
coordinates that need fitting.]

.Output doc file for plane parameters: planedoc001
[Enter output document file that will receive the three plane
 parameters. The equation for the plane is: z=ax+by+c. a,b, and
 c are determined.]

.Enter colums of x,y,z: 1,2,3
[Enter the columns where x y and z are to be found. (document file
 style counting)]

.Enter first, last key to be used: 7,12
[Enter the key range of the coordinates that are bein used.
 Restricting the range may be useful if not all keys have
 usable values.]

.Enter key for output file entry: 1
[Enter the key number where the results should be stored. This
 program enters only one line into the output document file. The
 key may be for example a loop index]
```

Programs: callfitplane.f, fitplane.f and support routines.
Author: M.Radermacher

fittomoexposure Fit a curve to the average densities of a file series.

file format: SPIDER document file

USAGE: fit tomo exposure or fittomoexposure

```
          .Input document file: indoc001
          [Enter the name of the document file containing the average values
           of the images in the tilt series. The document file must have the
           format: key 3 angle averages flag
           The flag should have the value 1. At this time it is not used but
           may be activated in the future to exclude single values.]

          .Output document file: fitdoc001
          [Enter name of output document file. This file has 5 values per line:
           angle  flag
           For correction, divide by the fitted polynome values.]

          .Degree of polynome): 2
          [Enter the degree of the polynome you want to use to correct the
           density values. At this time, only the value 2 has been tested,
           which should be sufficient for most cases. ]

          .Update factor in iteration: 0.5
          [The least suares fit is done iteratively. Often the algorithm is
           more stable when not the complete amount of corrections are applied
           in each iteration. This value must be smaller than 1. 0.5 is a good
           value.]

          .Print curves to results file (Y/N): N
          [Enter if you would like the curves of the fit in the results file.
           This is good for first checking, but creates a large amount of
           output.]
```

Programs: em_fittomoexposure.py, mircol.f, polyderiv.f, mrkur2.f, mrdocopen.f,
          mrdocread.f, mrmat3.f, fitfortomonorm.f (main).

Remark: mircol.f, mrkur2.f mrmat3.f and polyderiv.f are based on old programs
        developed in connection to my thesis work.

Author(s): M. Radermacher

for Syntax of for-loop

```
USAGE:        for v.variable from v.start to v.end step v.step do
        ... satements ...
        enddo

        The v.variable is the loop index
        The loop starts with v.variable = 1 and continues
        to v.variable = v.end.
        v.step is the step size of the loop and is optional.

        During the loop the statements between for and enddo
        are carried out.

        Examples:
        Copy 10 files:
        for v.index from 1 to 10 do
        copy
        file[3] v.i
        out[3] v.i
        enddo

        Copy every second file:
        for v.index from 1 to 10 step 2 do
        copy
        file[3] v.i
        out[3] v.i
        enddo

        Copy files with numbers specified in a list:
        [Note that the list index follows the Pyhton
         standards and the first index is 0 (therefore
         the last index of the list is 4)]
        v.list=[1,10,12,27,30]
        for v.i from 0 to 4 do
        v.num=v.list[v.i]
        copy
        file[3] v.num
        out[3] v.num
        enddo

Program: core section of EMIRA
Author: M. Radermacher
```

four Calculate the Fourier transform of an image of volume

file format: SPIDER

USAGE: four

.Input file: image001

.Output file: image002

REMARK: The Fourier transform works in both directions.

Programs: em-four.py, fourier_m.f and
subroutines. Programs estracted from SPIDER.

Author(s): M.Radermacher, others.

fourfilter **Apply a Fourier Filter to the Fourier transform** of an image

file format: SPIDER (fourier)

DESCRIPTION: This program can apply a variety of mostly low-pass and high-pass filters to 2D and 3D Fourier transforms, that comply with the default standard created by the command "four" (same format as SPIDER-fouriers).

USAGE: fourfilter

```
Fourier Input file: fourier001
[Enter input file name]

Fourier Output file: fourfilter001
[Enter name for filtered Fourier output file. To obtain the
filterd image the invers Fourier transform needs to be calculated
using "four"]

Filter options:
.  1, 2 cutoff low-pass and high-pass
.  3, 4 Gaussian low-pass and high pass
.  5, 6 Fermi low-pass and high pass
.  7, 8 Butterworth low-pass and highpass
.  9    Remez filter
. 10    B-factor

Filter type (1-10): 5
[Enter the filter type that you want to apply.]

For cutoff Gaussian and Fermi filter (type 1-6):
   Radius: 0.1
  [Enter radius in absolute Fourier units. If a radius larger than
  0.5 is entered the program will instead interpret this number as
  Fourier sampling units.]

For Fermi filter (5 and 6):
   Temperature: 0.02
   [Enter Fermi filter temperature factor. 0.02 gives a reasonably
   good ripple suppression. For weak highpass filters (small radius)
   the temperature can be adjusted to 0.01. Otherwise the low
   frequencies may not be sufficiently supressed. The value, however
   should be tuned to the application.]

For Butterworth filter (7 and 8):
    Passband, Stopband: 0.1,0.2
    [Enter the pass-band and stop-band of the filter. The order
    of the filter is derived from these values.]

 For Remez filter (9):
    Remez filter file name: remezfilter001
    [Enter the file that contains the remez filter values. For more
    detail see manual of "ff" command in spider.]

 For B-factor:
    Cutoff, B-factor, D-constant: 0.5, 1.0, 0.4
    [Enter the frequency cutoff, the B-factor and D, a constant to
    adjust the filtering.for more detail see "ff" command in the
    SPIDER manual]
```

REMARKS: This is a standalone version of the Fourier filter in SPIDER.

Programs: em_fourfilter.py, fourfilter_m.f, (adapted from four1a.f in spider.) ffilts_m.f, (adapted from ffilts.f in spider).

Author(s): M. Radermacher, 2012 (.py), 2008 (.f), prior authors: Many.

frc calculate Fourier Ring Correlation

image format: SPIDER

USAGE: FRC

```
        .Input file 1: image001
        [Enter first image for resolution determination.]

        .Input file 2: image002
        [Enter second image to be compared to the first.]

        .Output document file: resoldoc001
        [Enter name of output document file, receiving the following
         values:
         key,5,abs.radius,phaseresidual,FRC,Noise comparison,points in ring.
         Note: the phaseresidual value may be meaningless, see below.]

        .Ring width, Noise multiplier: 1,5
        [Enter the width of the ring to use for the calculation,
         and the noise multiplier that is used for comparison.
         (Noise comparison based on the number of points in a ring
         is not very meaningful though.)]
```

Programs: rfactsdo.f, rfacts2d.f
  Programs extracted from SPIDER.

Remark: These programs were extracted from SPIDER. However, the scaling of the phase residual does not work anymore (since ~1996) and we were sofar unable to fix it. In the latest version of spider the document file indicates the scaling failure (not in this version).

Author(s): M.Radermacher, others.

getfilenumbers **extract the file numbers from a file series and put** them in a document file.

file format: output: SPIDER document file

USAGE: extract filenumbers

.File pattern: image*
[Enter input image file pattern.]

.Output document file: filnumdoc001
[Enter the output document file, where the key will be numbered through followed by the file number in column 1.]

REMARK: This program was written to extract the file numbers from a series of micrographs. Only simple situations can be handled, where there is only one number inside the file. Mor complicated pattern require update of the python module.]

Programs: em_getfilenumbers.py (this is not a wrapper).

Author: M. Radermacher

getheader Get header information from SPIDER file

```
file format: SPIDER

USAGE:  getheader

        .Input file name: image001
        [Enter the name of the input file]

        .Results tuple or nothing: v.header
        [Enter a variable to receive the header information. If
         nothing is entered then the program uses v.|input file|,
         in this case v.image001. The header can be viewed using:
         print v.header. Single items can be retrieved for example
         as v.xdimension=v.header.nx
         The tuple items are:
        nx      x-dimenension                    (hdr[12])
        ny      y-dimension                     (hdr[2])
        nz      z-dimension                    (hdr[1])
        iform      file format                    (hdr[5])
        max      maximum (if known, else 0)     (hdr[7])
        min      minimum (if known, else 0)     (hdr[8])
        avg      average (if known, else 0)     (hdr[9])
        sig      sigma       (if known, else 0)      (hdr[10])
        phi      1. Euler angle around z        (hdr[15])
        theta      2. Euler angle around y'       (hdr[16])
        psi      3. Euler angle around z''      (hdr[17])
        xoff      shift vector x                     (hdr[18])
        yoff      shift vector y                    (hdr[19])
        zoff      shift vector z                     (hdr[20])
        stack      stack indicator                    (hdr[24])
        maxim=hdr[26]
        imgnum=hdr[27]
        lastindx=hdr[28]
        Second set of angles/shifts (the meaning of these
          is different in the later versions of SPIDER (after
          version 5.0., instead of second shift vectore, the
          header positions contain a third set of Euler angles.
          The Radon transform programs in EMIRA use the original
          definitions as defined below. A third angle position
          was abandoned since it is now occupied by the pixel size.
        phi2      1. Euler angle around z        (hdr[32])
        theta2      2. Euler angle around y'       (hdr[33])
        psi2      3. Euler angle around z''      (hdr[34])
        xoff2      shift vector x                     (hdr[35])
        yoff2      shift vector y                    (hdr[36])
        zoff2      shift vector z                     (hdr[37])
        pixsize      pixel size                     (hdr[38])
        ev      High voltage                      (hdr[39])
        Radon transform angles:
        rphifrom  starting angle 2D/3D Radon      (hdr[51])
        rphito       ending angle 2D/3D Radon           (hdr[52])
        rphiinc   phi increment 2D/3D Radon       (hdr[53])
        rthetafrom  theta start angle 3D Radon      (hdr[54])
        rthetato    theta end angle 3D Radon      (hdr[55])
        rthetainc   theta increment 3D Radon      (hdr[56])
        ]

Program: sp_header.py, adapted from SPIDER package.
```

getprocenters Program to get projection centers in a tilt series
(3 Euler angles) based on a 3D coordinate.

file format: SPIDER document file, (images optional)

USAGE: getprocenters

.Dimensions of the full tomogram, dx,dy,dz: 2048,2048,600
[Enter the dimensions of the full tomogram from which the coordinates were picked.]

.Center coordinate of the subtomogram, cx,cy,cz: 500,500,101
[Enter the x-y center coordinates as found in the large 3D reconstruction. The z-center coordinate, however, is calculated in the smaller subtomogram. For example in a sub tomogram of z-dimension 200, the value 101 for z would put the center in the standard center (SPIDER convention). To convert the z coordinate picked in imod that then is applied to SPIDER converted projectionsm cz needs to be calculated as: zs/2-zf/2+z, where zs is the -dimension of the small tomogram, zf is the z-dimension of the large tomogram and z is the z-coordinate picke in the large tomogram.
If 0 is answered cx,cy will be the center of the 0-degree projection of the large reconstruction and cz the z-center of the new reconstruction. i.e. the 3D structure will not be moved in z-direcation.]

.Enlargment of tomogram: 1.0
[if there is an enlargment between the projections and the 3D reconstructions, specify it here.]

.First, last projection number: 1,92
[Enter the first and last projection number that you want to process.]

.Read projection angles from doc file [Y/N]: Y

If Y is answered: (NOTE: THIS OPTION SHOULD WORK BUT NEEDS MORE TESTING.
Please let me know about good and bad experiences with it.)

.Input doc file with projection angles: projdoc001
[Enter the document file that contains the projection angles.]

.Doc file columns containing phi, theta, psi: 1,2,3
[Specify the comlumns in the document file that contain the angles.]

.Dimensions of large projections: 2048,2048
[Enter the dimensions of the large projections.]

if N was answererd:

.Name of first projection: proj001
[Enter the name of the first projection.

Programs: getprocenters.f, calcprocenter
AUTHOR: M.Radermacher

getvalue Get shift vector from pair of alignment files or single value

file format: SPIDER

USAGE: get value

```
        Input file: shift001
        [Enter name of 3D file that contains the shift values obtained
         by rad2d3dali. Alternatively any image/volume can be used here
         to get a single pixel/voxel value.]

        Direction file or *: angle180
        [Enter the name of the 3D file containing the angles for the
         shift direction obtained by rad2d3dali. If * is entered only
         the value from the first file will be returned.]

        Coordinates ix,iy,iz: 33,33,1
        [Enter the coordinates of the point whose values should be returned.]

        Enter variable to receive output :v.value
        [Enter the name for the named tuple to receive the output. The return
        format will be:
        getimgvalue(value=1.0, angle=180.0, xshift=-1.0, yshift=-3.2584e-07)
        and acessible for example through v.value.xshift.]
```

Remark: While the program was mainly written to extract the projection shift vector from the alignment files, if can be used to just get a single value froma file. The tuple then has only a value in the first place and all else is 0.

Program: getvalue.f, em_getvalue.py

Author: M. Radermacher

hsearch lorentz Search for helical pitch

file format: SPIDER

USAGE: hsearch lorentz (or hsearchlorentz)

```
        .Input 3D file: helix001
        [ Enter input image or volume]

        .Output symdoc file: symdoc001
        [Enter name of output document file, which will, after each run
        contain the last best match information. ]

        .Pixel size, min, max radius: 3.06, 0, 60.
        [Enter the pixel size in Angstroems and the minimum andmaximum
         radius.]

        Increments in phi,z : 1.0,1.
        [Enter the search increments for phi anz z]
```

REMARK: This module requires the presence of Ed Egelman's helix package (V1.0)

Programs: em_hsearchlorentz.py (in directory egelprogs)
  hsearch_lorentz_rss.F

Author(s): M.Radermacher (python), Ed Egelmen

Requires: Installation of Ed Egelman's helix package

if if-construct

Purpose: conditional statement in batch/script file.

Example: if(v.x > 2.) then
         statements.....
         endif

         The if statements replicates the if of python.
         The comparison operators are:
         <   less than
         <=   less or equal
         >   greater than
         >=   greater or equal
         !=   not equal
         ==   equal
         The string comparisons as "in" and "not in" or "is"
         have not been tested but may work.

Note: Currently there is no 'else' but may be created later.
      It is not planned to introduce an elseif since this would
      make the syntax unnecessarily complicated.

imcfromppcaem Convert the ppcaem output coordinate file to
an IMC-file compatible with spider.

file format: see below

USAGE: imc from ppcaem

.Input coordinate file from ppcaem: coords001
[Enter the coordinate file you got from ppcaem.]

.Output IMC-file: IMC001 or PREF_IMC
[Enter the name of the spider compatible image coordinate file.
for SPIDER 5.0 this must be IMC followed by 3 digits, like IMC001,
for 2014 spider (and, unfortunatley for the EMIRA version of the
extracted programs) the file name must be a prefix followed by
inderline followed by IMC. No number at the end. This is the only
file type in spider that does not have at least a 3 digit number.
If any numbering is desired you can include a number in the prefix.
In emira you then also could loop over it.]

.Number of columns to read: 12
[Ente the number of columns to read in the ppcaem file. The maximum
line length the program reads at the moment is 180 characters. If
the line is longer you will get an error. Fix requires a change in
dimensions in imcfromppcaem.f.]

.(O)ld or (N)ew format: N
[See remark above. The file name you gave should be consistent with
the format, otherwise the spider programs that follow correspondence
analysis cannot handle it.]

.Scaling factor for coordinates: 0.1
[Enter a scaling factor. During the test, the corrdinates were very
lage (10**6) and it seemed approrpiate to make the smaller for more
stable calculations.]

.Xdim, Ydim if known: 0,0
[The spider IMC file stores the original x-y dimensions of the input
images. There is no space or a z-dimension. The dimensions are not
in the ppcaem coordinate file. The default of 100,100 is used if 0 or
nothing is entered.]

Program: imcfromppcaem.f
Author: M. Radermacher, 2014

imodcoordinates2doc Convert 3D coordinate file from IMOD to SPIDER-style document file.

file format: see below

USAGE: imodcoordinates2doc

.Input imod coordinates file: imodcoos001.txt
[Enter the coordinate file produced by IMOD. This file must be created with model2point -float]

.Output doc file: coodoc001
[Enter the name for the spider-format document file]

.Number of columns in input (3 or 4): 4
[Enter the number of columns in the input file. Depending on the option used in the model2point command in IMOD the file will have 3 or 4 columns]

Programs: em-imodcoordinates2doc.py, /extras/fstandalone/converters: convert_imod_spider.f

Author(s): M.Radermacher

mask Apply a round mask to an image or volume

file format: SPIDER

USAGE: mask

```
        .Input file: image001
        [ Enter input image or volume]

        .Output file: subdir/immasked001
        [Enter name of output file. If the directory does not
        exist, it is created].

        .Outer, inner radius: 30,5
        [Enter an outer and an inner radius. (defaults: Outer radius
        20000., inner radius 0.) ]

        .Center of mask, x,y,z: 33,33,33
        [Enter the location of the mask center.]

        .Edge type (Disk,Cosine,Gaussian,True Gaussian): D
        [Enter the first letter of the shape options.]

        For everything except option D:
            .Falloff in pixels: 7
            [Enter the falloff of the mask. Default=5.]

        .Background type Average,Prec.Average,Circumf,Extern:E
        [Enter background option. Average = average of image before
        masking, Precise Average=average of density within the mask.

    If E was answered:
        Bachground value: 0.
        [Enter background value for the image parts masked out.


        .Enter variable to receive output: v.stats
        [The output variable will receive a named tuple. To look
        at it use the command print, to use the results acces the
        return values as: v.stats.min, v.stats.max, v.stats.avg and
        v.stats.sig. These can be used like normal vaiables. If no
        answer was given for the output variable, then the name
        v.filestats is used by default.]
```

Programs: mask_m.f, mask.f, Program estracted from SPIDER.

Author(s): M.Radermacher, others.

mode **set some behaviors of EMIRA**

```
USAGE:        mode

        The program will ist the available modes. At this time they are:

        display on   -- activate auto display (output images only)
                This is on by default for interactive, singe command use.
                Needs to be switched on if desired in batch mode.

        display inout  -- activate auto display (input and output images)
                Not active by default.

        display off  -- deactivate auto display
                Default in batch mode.

        .Set mode: display off
        [Enter mode.]

Program: em_mode.py
```

model Create model image (2D)

```
image format: SPIDER

USAGE:  Model

   .Output file: modelimage001
     [Enter name of output image.]

   .x,y dimensions (def 8,8): 128,128
     [Enter the dimensions of the output image.]

   .Options for models are (2 first letters suffice):
   .  Blank    -- blank image (is the default)
   .  Disk   -- filled cirlce (disk)
   .  Ring     -- Ring
   .  Gaussian -- Gaussian function
   .  Random   -- Random distribution
   .  Sine     -- set of sine waves
   .  Test     -- 2D sine wave
   .  Wedge    -- Density wedge

   .Enter type: blank
     [Enter which type of model image you want to create.]

 if blank was answered:
   .Backgound (def. 0): 1
   [This would create an image with constant value 1.]

 if ring or disk was answered:

   .Outer radius: 60
   [Enter the outer radius if the disk. will be filled with the
    value  1, outside 0.]            .

 if ring was answered:
   .Inner radius: 10
   [Enter the inner radius of the ring.]

 if Gaussian was answered:
   .Center coordinates x,y: 45,45
   [Enter the center coordinates of the Gaussian density.]

   .Sigma in x,y: 5,5
   [Enter the sigma of the Gaussian density. If the sigma in
    x is different from the sigma in y then the distribution
    will be eliptical in shape.]

  if random was requested:
    An image with random distributed densities will be created.
    .True Gaussian (Y/N): Y
    [If Y was answered the distribution will be Gaussion, otherwise
     just what the random generator creates.]

    if true Gaussion:
           .mean, sigma (def: 0,1): 0,1
           [Enter the mean and the sigma of the Gaussion distribution.

  if sine wave is requested:
    .Amplitude and phase (def: 1,0.): 1.,0.
    [Enter amplitude and phase of the sine wave.]

    .Frequency in x,y: 1,0
    [Enter the frequency, calculated as periods/image.]

  if Test of Wedge is requested, no additional answer is needed.

Program: model.f extracted from SPIDER.
Author: M.Radermacher of em_model.py, many authors for model.f
```

montage montage many images into one large one

file format: SPIDER

USAGE: montage

```
        .Input image with lowest number: image001
        [Enter an example image file name with the smallest image number.
        The important part is that the number of digits >0 in the example
        image is not larger that the number of digits in the image with
        the lowest number that is being used. The file number is replaced
        counting from the back, i.e. the number 12 would replace only the
        last two digits, and leave everything in fromt of it as is.]

        .Image numbers: 1,3,5,7-10,12
        [Enter the image numbers. This can be a comma separated list or,
        like above a from-to answer, or a mixture of both. Must not
        exceed one line.]

        .Images per row, image margin: 4,2
        [Enter the number of images per row, and the margin width.]

        .Margin value: 1.5
        [Enter the density value to be place in the margin areas.]

        .Normalize to 2 (Y/N): Y
        [Answer if all individual images are normalized to a range
        of 0 to 2.]

        .Output file: mont001
        [Enter the name for the output image that will contain the montage.
        If a directory is specified that does not exist, it will be created.]
```

Programs: montage.f, in /extras/fstandalone/applications. Program
  estracted from SPIDER.

Author(s): M.Radermacher, others.

MSA Run multivariate statistical analysis

file format: SPIDER

SEE ALSO: msaimc2doc, msamap, msavismap

USAGE: msa

```
          .Input file example: image001
          [Enter the name of the image or volume containing densities]

          .Use number string:N, or document file:D : N

           If N was answered:

               .File numbers: 1-100,210-200,401-10000
             [Enter string that describes file numbers. Program cannot handle
              missing files (the old coran in SPIDER 5.0 could do this). So, if
              you have gaps, which is common, use the document file option. ]

           If D was answered:

                 Document file name with file numbers: Filelistdoc001
              [Enter the name of a document file that has the file numbers or a 0/1
               selection file. Both should work, see question below.]

              Column number to use (0 = use as 0/1 selection file): 3
              [Enter the column number that contains the file number. If 0
              is entered then the program uses column 1 as a 0/1 indicating if
              the file is to be used, and the key column as file number.]

           .Mask file: mask001
           [Enter the name of the file containing the image binary image mask.
            Only the pixels in the image that coinside with the 1s in the mask
            are used for MSA analysis.]

           .Number of factors: 8
           [Enter the number of factors (eigenvectors) to be calculated.]

           .Run CORAN (def.), or PCA: Coran
           [Enter which type of analysis to run. Only the first letter is
           used, but the answer may be longer.]

           .Additive constant: 1.
           [Enter an additive constant so that for correspondence analysis
           there are no negative vlaues in the data.]

           .Prefix for output files: molecule
           [Enter a prefic that is used to construct several file names. These
           are: _IMC., _PIX., and _EIG..
           Unfortunately these file names do not comply with spider file
           naming conventions. However, later programs need them also with
           this type of names. (this program was extracted from spider).
```

Notes: The program was extracted from spider.

Programs: em_msa.py, jpmsk1.f and subrountines

Author(s): M. Radermacher: for .py and rewrite of jpmsk1 for stanalone,
J.P. Bretaudier for the original JPMSK1 and subroutines, and
members of the Albany lab for the rewrite of a large part of the
code.

msa imc2doc Convert image coordinate file created by MSA to document file

file format: input: special, output: SPIDER document file

USAGE: msa imc2doc

.Input IMC file: Prefix_IMC
[Enter coordinate file from msa (coran or PCA). This follows the unfortunate post 2008 SPIDER convention, where the MSA files do not have a number anymore, but only a prefix.]

.Output file: imcdoc001
[Enter name of output document file. The document file will have the image number as a key, followed by the coordinates.]

PROGRAMS: imc2doc.f

Author(s): M. Radermacher

msamap **Print 2D maps from PCA or coran coordinates.**

file format: special

USAGE: msa map

.I for image, P for Pixel: I
[Enter if you want a image corrdinate of a Pixel coordinate plot.]

.Prefix of input coordinate files: three
[Enter the prefix of the coordinate files. The anser "three" will create the filename "three_IMC" for example.]

.Number of horizontal patches: 0
[Enter the number of horizontal patches, if this is a correpondence analysis for patch averaging. Options not 0 have not been tested in this version.]

.Two factor numbers: 1,2
[Enter which 2D map you want to print. The above anser would print map 1 (horizontal) versus 2 (vertical).

.S symbol, I id: s
[Enter if you want to print the id-numbers on the map or use symbols. id-numbers allow to pinpoint single images, but because the numbers take a lot of space, that may not be recognizable. Symbols, well choses, make it easy to observe the coordinate pattern, but don't allow for an easy identification of each image.]

if "S" was ansered:
.Enter symbol: +
[Enter the symbol you want to print at each coordinate. A "." uis often the best.]

.Postscript output Y/N: Y
[Enter if you like to have the output in a postscript file. Otherwise the maps are ptinted in the Results file.]

If Y was answered:
.Enter name for postscript file: mappost012
[Enter the name of the output file. A good nameing convention is to use he file number for specifying the map content. For example mappost045 would be map 4 versus 5. The first digit then can be used if several versions are created etc.]

If N was answered:
.Number or pages, Number of line: 0,0
[You can specify the number of pages a map should be large and the number of lines. By default the map fits on a page with 132 character width.]

.Std. deviation: 5.0
[Enter the standard deviation, above which the map wil lbe cropped. The default is 10, essentially no cropping. For first runs, this valus should be large, for example to identify outliers, but also to identify small groups of particles which may form a class that would get lost.]

.Flip options 1,2, or 4, flip f1, flip f2 or both: 0
[If you want to flip one or both of the axis, spacify this here.]

Notes:

Programs: sgraf.f, (subroutines in file).

Author(s): M. Radermacher, J. Frank, and others.

msamaptoimage create an image containing an MSA image coordinate map

```
file format: SPIDER

USAGE:       msa map to image

        .Input coordinate file (full name): prefix_IMC
        [ Enter the name of an image coordinate file created by one of the
          MSA programs.]

        .Output image file: mapimage001
        [Enter name of output file that will contain the map.]

        .Output document file name: imagedoc001
        [Enter name for the document file that will contain:
        key, image number, x-coordinate, y-coordinate
        inside the crated imag. The purpose of the file is to make
        it possible to find an image number from the coordinates.]

        .Sigma factor for map (0 if no limit): 5.0
        [Enter a limit for the map. The map will be cutoff at x-average
        +- sigma*signamfacor, y-average +- sigma*sigmafactor. The document
        file will also contain the coordinates outside of the image range.]

        .Two factor numbers: 1,2
        [Enter the factor number that will span the map.]

        .Image dimensions (x,y): 1000,1000
        [Enter the dimensions of the image. The coordinates are scaled to fit
         in the image.]

Author:         M. Radermacher
Programs: imagemap.f
```

msamaptoline Projects the points from a 2D MSA map onto lines

```
image format: SPIDER, IMC file

USAGE:    msa map to line

          .Input coordinate file (full name): prefix_IMC
          [Enter the name of the coordinate file obtained by corrrepondence
           analysis, Principal component, non-linear mapping or converted
           from PPCAEM.]

          .Options: P
          [There are sevaral options.
           Option A: Specify an angle of the projection
           Option P: Specify 2 points that define a line perpendicular to
                       the projection (The projection is caclulated in the
                     direction of this line.).
           Option R: Calculate toe projections at all angles (essentially
                     a Radon transform of the map.)

           .Name for line projection document file: profile001
           [Enter the name for a document file containing the line projection
            in histogram form.
            if option R is used this is the name of the output Radon transform.]

           If R was NOT specified:
           .Name for sorted images document file: sortdoc001
           [Enter the name for a document file that contains the image numbers
            sorted as they are projected onto a line. If the map shows two
            main clusters and is projected in a direction that separates the
            cluster, this document file can be used to separate the data
            accordingly.]

           .Dimension (length) of line: 200
           [For simple line projection this specifies the number of entries in
            the document file and thus the histogram divisions. For option R
            it defines the p-dimension of the Radon transform.]

           .Two factor numbers for map: 1,2
           [Specify the map that should be projected.]

           if option P was used specify two point:
           .Enter point coordinates x1,y1,x2,y2: 10,10,200,200
           [Enter the two points that determine the direction of the projection.
            in the example entered this would be a 45 Deg.]

           Without option P:
           .Enter line angle or increment: 10
           [Enter the angle of the projection. If option R was specified this
            value is used as the increment for the Radon transform.]

Remark: This is a first program to separate two clusters in a map. More
        complicated separation algorithm will follow.

Author:    M. Radermacher
Programs:  maplineproject.f

           Enter if this is a new 3D to be created (Y), or an existing 3D that
           you want to add to (N or any other letter).]

          If it is a new 3D:
             .x,y,z-dimension: 2048,2048,800
             [Enter the dimension of the volume to create.]

          .ix,iy,iz center coordinate: 0,0,0
          [Enter the coordinates for the center if you are calculating
           a subtomogram.
           The x,y coordinates are relative to the 0 degree
           projection, using the conventions of "pixel" in SPIDER WEB. I.e.
           x and y start at array coordinates (1,1).
           z is the location of the center of the volume. If 0,0,0 is answered
           the origin of the 3D is placed at te default center:
           (x-dim/1+1, y-dim/2+1, z-dim/2+1.)]. If z was determined in imod,
```

then the answer here should be: Ds/2-Do/2+Zm, where Ds is the
z-dimensions of the small volume, Do is the z-dimension of the
volume from which the coordinate was picked, Zm is the picked
z-coordinate.

.Enlargement: 1.0
[Enlargement allows for different magnifications in 3D and 2D. It
also may be smaller than 1 to create a size reduced reconstrution.
Note however, that demagnification may create a noisier
reconstruction because of possible subsampling.]

.Name of first projection: proj001
[Enter the name of the first projection in the series. The
projections must have the projection angles in the header in header
position 1. They may contain all 3 Euler angles.]

.First, last projection number: 1,72
[Enter the first and last projection number.]

an optional scaling factor for the density values.]

Notes: The fortran programs are extensions of the original backprojection written 1978.(see M. Radermacher, W. Hoppe, 3-D Reconstruction from Conically tilted Projections. Proc. 9th Int. Congr. El. Micr. Toronto 1978, Vol I 218-219 and M. Radermacher, thesis, 1980). For each projection the algorithm loops through the volume and adds the interpolated projection value to the 3D volume. Modifications include the extension to 3 Euler angles (instead of 2 angles in the original), and the allocation of the volume in-core (instead of one slice at a time).

For large volume make sure that the memory of your computer is sufficiently large to avoid usage of swap space. Rule of thumb, the available memory should be at least 1.3 times the size of the volume. If memory is insufficient the volume should be reconstructed in chunks, which can be done by using the subvolume options.

Programs: em_backproject.py, bapi3absubv.f, mirbp3absub.f

Author(s): M. Radermacher

msanlm Apply nonlinear mapping to the image coordinates after msa.

file format: specialformats for image coordinates (SPIDER new MSA compliant)
SPIDER document file format

SEE ALSO: msa, msaimc2doc, msamap, msavismap

USAGE: msa nlm

.Full name of image coordinate (IMC) file: nlmin_IMC
[Enter the file containing the image coordinates. The file has
the format the msa creates. Essential is the following:
Line 1 integers,
<# of coordinates>, <# of factors>, , ,
<# of coordinates(again)>, <0>.
The image dimensions are not needed for nlm.
following lines contain image coordinates, folowed by two more
numbers, followed by image id number (float), followed by "1.00"
The file is formatted with up to 10 real numbers per line. More
details may be avaliable in the SPIDER description of PCA/CA]

.factor number to use: 1-6
[Enter which factors you like to use for the nonlinear map.]

.Doc-file for 2D output map: nlmdoc002
[Document file containing the 2D output coordinates. Key =
image identifier number.]

.Name for IMC style 2D output: nlmout002_IMC
[Will contain the same info as the document file, but in a format
that can be further evaluated with programs to analyze msa maps,
for example visual maps. It has the same format as the output of
MSA (and the input to this program, above).]

.Name for postscript map file: nlmpost002
[Postscript file with plot of image numbers. For large sets this
may not be very readable.]

.Map standard deviation limit: 10
[set a limit to the map display in the postscript file.]

.Enter W to apply weights: n
[You can apply a renormalization to the factors. At this time the
weights are normalized with 1/(5sigma). This will need future
revision and tuning.]

.Number of iterations: 1000
[Enter the maximum number of iterations. They will be reached if
the epsilon (below) is never achieved. If in the last step though
the errorsome increased that the iterations will continue until at
least the last lowest error is achieved within 0.1 per mille.]

.Magic factor for steepest desc. algor.: 0.4
[Enter the "magic factor" used in the steppest descent iterations.
The default, and value recommended in the literature is 0.4.
However, it seems that when the dimensional reduction is very
large, smaller magic factors should be advantageous (e.g. 0.2 or 0.1).
(this will be further tested and additional recommendations will come.)]

.Epsilon for iteration cutoff: 0.00001
[Terminate the iterations if the error is below this value. Check this
value. In many situation even 0.1 is a rather stringent limit.]

.Exponent in error measure: 0.0
[Determined the type of optimization. If the exponent is 0 then
the long distances will get highest weights (Kruskal,
Psycometrica, 29, 1964, 1 and 115). If it is 1 than the short distances
are more important (Sammon, IEEE Trans. Computers, C-18 (1969), p. 401).
In-between values may be used.]

```
        .Lower distance threshold: 0.0
         [ you can specify a threshold below which coordinates are considered
         identical and removed from the calculations, retaining only one of
         them.]

        .Start distribution type: 1
         [Enter the type of start distributions. Currently implemented:
         1 -> specify a coordinate pair
         2 -> Random distribution
         4 -> read start distribution from file

        if 1 was answered:
        .Factors for start distribution: 1,2
        [Enter the factor pair to be used for the starting map.]

       if 4 was answered:
        .Document file with 2D image map: nlmdoc001

Notes:  The program was extracted from spider where it had become
         obsolete in the transitions from VMS to UNIX. The program
        here is a rewrite with substantial changes for better effciency
        and larger data sets.

Programs: em_msanlm.py,mrerr.f,disttst.f ,strtdis.f,distlst.f
           nonlmap.f,dist2d.f,mrnlstrt.f (main program),nlmrealstrt.f
           nltfile.f, mrdev.f

Remark: Advise is to first optimize the large distance in a coordinate
        set and then optimize the short disances, using the large distance map
        as start distribution. It can be shown that this is similar to
        a manifold mapping if used in this order. (Citation needs to be looked
        up).

Author(s): M. Radermacher. See Use of Nonlinear Mapping in Multivariate Image
Analysis of Molecule Projections, Ultramicroscopy `17 (1985) 117-126. (see also
Erratum since a division in the equations in the paper is printed wrong:
Ultramicroscopy, 19(1986), p75)
```

msaselectbymask select images according to mask of map

File formats: SPIDER image, and document files

PURPOSE: Select a set of image coordinates according to a mask applied
to an image displaying the images as points. (see msa map to image).

USAGE; msa select by mask

.Mask input file: mask001
[Enter the image file containing the mask for the map.]

.Input document file: imagedoc001
[Enter the name of the document file that contains the coordinates
of the images scaled to the image map.]

.Output document file: clusterdoc001
[Enter the name of the output document file that will contain the
image numbers and coordinates of the images tha fall under the
mask.]

.Position of: image #, x and y: 1,2,3
[Enter column numbers of the image number, the X-coordinates and the
y-coordinates in the input document file. This file should have
been created by "map to image".]

.Select above or below 0.5 (a/b): b
[Enter if the selection is for all mask value below 0.5 or above 0.5.
The reason for this choice is that a hand drawn mask in WEB may be
of value 0 in a background of 1.]

Programs: extractcoos.f
Author: M. Radermacher, 2016

msavismap Create visual representation of PCA map

File formats: SPIDER

```
        .Output file: vismap001
        [Enter the name of the output file that will contain the map
         as an image.]

        .Map dimensions (x,y): 1200,1200
        [Enter the sice of the image to contain the map. It is avisable to
        kepp the size such that it can be displayed on the monitor screen.]

        ,Number of divisions x,y: 10,10
        [Define the grid of the map. Specify the disvisions in x and y
         direction. When the map is assembled, all images, whose coordinates
         fit in the same gridsquare are averaged. The average is placed in
         the map image at the proper location. If the image dimensions
         differ from the space available within a grid square, the average
         image is either padded or cropped.]

        .Upper lower contrast limt in sigma units: 5,5
        [Contrast limit in multiple of sigmas, applied to the average images.
         Use generous values. This is mostly intended to elimiate dust and
         other extreme density values. If the threshold is set too small
         the images tend to look more binary, and it is possible to
         miss the most important features that led to their arrangment.]

        .Map output doc. file: vismapdoc001
        [This document file will contain the file numbers and the map square
         where they contribute.Format" key, image number, grid square number.
        grid square numbers are counted left tp right top to bottom. With a map
        division of 10x10 the grid numbres go from 1 to 100.]

        .Input img. coord. doc file: imcdoc001
        [Enter doc file with image coordinates. This must be already converted
        from any PCA/CA specific coordinate format. Use for example the command
        "msa imc2doc".

        .First key, last key: 1,23756
        [Enter the fist and last key in the document file. The range can
        be larger, and non-existand images will be skipped lateron.]

        . 2 Factor numbers (=cols in docfile): 1,2
        [Specify factor numbers for map. Here in the example the map
        eigenvector 1 vs eignevecto 2 is requested.]

        .Map standard deviation limit: 10
        [Specify the map standrd deviation. This means that all coordinates
        / images outside the range of limit*sigma are discarded from the
        visual map. Be generous with this value. If it is set too low,
        you are focusing only on the center of your map, maybe missing
        interesting features away from the center. If the limit is set larger
        then the range of the data, the value has no influence and the
        images still fill the map (i.e. it does not change increase the
        coordinate range of the map.]

        ,First input file: ../../alif/alif000001
        [Specify the name of the first imput image to be used in the process.
         Numbers are substituted by the numbers found in the coordinate
         document file.]
```

Program: em+vismap.py, vismap.f
Author: M.Radermacher (vismap was originally written for SPIDER)

multiply Multiply two images

file format: SPIDER

USAGE: multiply

```
        .First input file: image001
        [ Enter input image or volume]

        .Second input file: image002
        [Enter second input file.]

        .Output file: product001
        [Enter the name the output file receiving the product of the two
         input files.]
```

Programs: add.f, and subroutines
  estracted from SPIDER.

Author(s): M.Radermacher, others.

multiplyconjugate Multiply two Fourier transforms

file format: SPIDER

USAGE: multiply conjugate

```
        .Complex valued input file 1: four001
        [ Enter input file containing complex values]

        .Complex valued input file 2: four002
        [Enter second input file containing complex values. The conjugate
         of the second file will be multiplied with the first file]

        .Output file: product001
        [Enter the name the output file receiving the product of the two
         input files.]
```

Programs: add.f, option -2, and subroutines
  estracted from SPIDER.

Author(s): M.Radermacher, others.

new2olddoc Convert new format document file to old format document file,
readable by spider 5.0

file format: SPIDER document file

USAGE: new2olddoc

.Input new format doc file: newdoc001
[Enter the name of the document file you like to convert.]

.Output old forma doc file: olddoc001
[Enter the name for the spider old-format document file]

Programs: em_new2olddoc.py, /extras/fstandalone/converters:
new2olddoc.f

Author(s): M.Radermacher

oldimc2doc Convert IMC image coordinate file from older coran to doc file

file format: see below

USAGE: oldimc2doc

.Input imod coordinates file: IMC001
[Enter the coordinate file produced by CA S in SPIDER version 5.
(It is not clear with which spider version the format of this
file changed.)]

.Output doc file: imcdoc001
[Enter the name for the spider-format document file]

.Number of factors: 12
[Enter the number of factors in the input file that are to be
place into the output document file.]

Programs: em-oldimc2doc.py, oldimc2doc.f

Author(s): M.Radermacher

pba3 Peak file averaging for projection based 3D alignment

file format: SPIDER

See also: rad alignto3d

CAUTION, THE REWRITTEN PROGRAM IS NOT FULLY TESTED YET AND MAY NOT ALWAYS WORK.EXAMINE THE RESULTS CLOSELY BEFORE TRUSTING THEM.FURTHER DEVELOPMENT IS IN PROGRESS.

USAGE: PBA3

.First input peak file: peak001
[Enter the first peak file (3D CCF), from a reference projection alignment to the 3D volume of interest.]

.file numbers: 1-25
[Enter the file numbers of the peak files (equal to the projection file numbers).]

.Enter peak file number foe header info: 8
[Enter the number of the peak file that should be used for extracting the header info, which includes the angular increments.]

.Full range peak file Y/N: Y
[Enter Y if the input peak files cover the full range of the Euler angles. Note that this would be phi over 360 deg, theta from 0 to 180 deg (not -90 to 90 deg!), psi over 360 deg.]

If N was answered, i.e. the peakfiles are from a subsearch then:

.Enter center phi, theta, psi and range: 45,70,120,90
[Enter information to create the output peak file. Center coordinates phi, theta, psi, and the angular range (total) for the file. In the above example the range would be -45 to + 45 arond the center for all 3 angles. At this time it is not possible to give different ranges for the different angles.]

.Output averaged peak file: ccf3daverage001
[Enter the name of the output file that will contain the average.]

.Output counter file: outcount001
[Enter a name for the counter file. This is a file that keeps track of how many measurements were averaged at each angle. Note: for a full search the values should be the same in every point, and if you try to display it in WEB it will complain about blank image, which in reality may be simply a constant image.]

.output document file: 3dalidoc001
[Enter the name of the document file that will receive the 3D alignment values.]

.Enter key to store results: 1
[Enter the key for the line in the document file that will receive the results. In a loop over the alignment for several volumes this should typically be the volume number.]

.Input document file (or *): indoc001
[Enter the name of the document file that contains the angles of the projections that were used for the 3D projection alignment (the process that creates the peak files). The key should correspond to the peak file number and the columns behind should contain the three Euler angles of the projection.]

If a document file was provided then:

.Enter columns for phi, theta and psi: 1,2,3
[Enter the document file columns were phi theta and psi are found.]

If not document file is provided (answer was "*") then:

.First projection file: pro001
[Enter the name of the first projection file. The program will look up

```
         the projection angles in the headers of the projections. The numbering
         must be the same as the numbering of the peak files.]

        .Projection file numbers: 1-15
        [Enter the file numbers of the projections that belong to the peak
         files that are being averaged. Most of the time they are the same
         numbers, but not always. For example the peak files may be numbered
         301,302,303 etc, while the projections may be numbered 001,002,003.]

        .Header position of angles (1 or 2): 1
        [Enter if the angles are stored in angle position 1 or 2. (3 is not
         used anymore since the SPIDER header was changed and does not leave
         enough space for the complete alignment parameters anymore in three
         positions.]

        .Debug: N
        [Enter N. Only if you have a small job and really want to debug the
         program answer Y.]

Author: M. Radermacher. Programs: rmpeakavg2.f, combeul2.f, rangeangles2.f,
        em_pba3.py

Notes:  This is a FORTRAN redesign of pba3.c originally written by Lingbo Yu.
        This program does not take care of 3D shifts.
```

peak3 Find the maximum in a 3D file.

file format: spider

USAGE: peak 3

```
        .3D input file: peakfile001
        [Enter the name of the 3D file.]


        .Output document file or * : doc001
        [Enter the name of a document file that will receive the peak
         coordintaes and values.]

        .find maxima (+) or minima(-): +
        [Enter + if you are looking for maxima, - if you are looking for
         minima.]

        .Number of peaks to find: 5
        [Emter the number of peaks to look for.]

        .Redfine origin (Y/N): N
    If 'Y' was answered to origin redefinition:
                .New origin coordinates (x,y,z): 11,12,13
                [Enter new coorintes values for the origin of the
                 coordinate system.]



        .Restrict to a selected box: N
        [Enter if the search should be restructed to a box smaller than the
         complete volumes.]
    If box selection is chosen answer:
                .lower (x,y,z): 5,5,5
                .upper (x,y,z): 20,10,10
                [Enter the lower and upper value of the coordinates of the
                 selected box, where the maximum is  found]


        .Restrict search neighbors (Y/N): N
        [Answer Y if the search is restricted such, that peaks must be
         separated by a minimum distance.]
    If neighborhood restriction was chosen:
                .Neib. exclusion elipse axes (x,y,z): 5,4,5
                [Enter the axes of the eliipse surrounding a maximum, that
                 is excluded from the search.]

       Except when neighbourhood restriction is specified:
         .Determine center of gravity (Y/N): N
         [Enter 'Y' if the center of gravity of the peak is determined. By
          default, a quadratic interpolation of the peak coordinates is
          avalaible, which will be replaced by the center of gravity. This
          option is mutually exclusive to the neigborhood exclusion option.]


          Enter variable to receive output : v.peak
          [Enter the variable that will receive the coordinates of the
           highest peak and its value. The elements of this tuple will be:
           v.peak.nx - x-coordinate of peak counting from 1
           v.peak.ny - y-coordinate of peak counting from 1
           v.peak.nz - z-coordinate of peak counting from 1
           v.peak.cnx - x-coordinate of peak relative to volume center
           v.peak.cny - y-coordinate of peak relative to volume center
           v.peak.cnz - z-coordinate of peak relative to volume center
           v.peak.x - interpolated x-coordinate of peak relative to vol. center
           v.peak.y - interpolated y-coordinate of peak relative to vol. center
           v.peak.z - interpolated z-coordinate of peak relative to vol. center
           v.peak.value - value of maximum
           REMARK: The interpolated value is either a bilinear interpolated
           value of the center of gravity coordinates, if this was specified.
```

Programs: em_peak3.py, speak3.f pksr3.f, cgr3.f, rdsl_p.f

(mostly extracted and modified from SPIDER.)

Author: M. Radermacher and others

pickslice pick a slice from a volume

image format: SPIDER

USAGE: pickslice

```
        .Input file: volume001
        [Enter input volume name.]

        .Output file: slice001
        [Enter output image name. Directory will be created if needed.]

        .slice direction (def. Y): Z
        [Enter the slicing axis.]

        .slice number: 101
        [Enter the slice number (coordinate along slice direction).]
```

Programs: pickslice.f,picksl.f,picksv.f and subroutines.
  Some programs extracted from SPIDER.

Author(s): M.Radermacher, others.

pickstickturn Program to box out the projection series of an
elongated object (stick) from a tomogram, and
update the tilt angles such that the reconstruction
will show the object orientated along the y-axis.

file format: SPIDER document file

USAGE: pickstickturn

.Input doc file with picked coordinates: windoc015
[Enter the name of the document file that contains information
about the subvolume and the tomogram itself, in total 12 values
per line.

The format is:
N 12 x1,y1,z1,x2,y2,z2,xdim,ydim,zdim,astart,n1,nlast
with:
x1,y1,z1 Coordinates of first point
x2,y2,z2 Coordinates of second point
xdim,ydim,zdim Dimensions of the large volume
astart Start angle used during tiltseries recording
n1 Filenumber of first projection in the series
nlast File number of last projection the tiltseries.
(n1 and nlast are used as keys in the output file)
]

.Prefix for projection-set docfiles: prodoc
[Enter a prefix for the document files that contain the projection
information for every subvolume. This file will contain one
line per projection and have the following information:

N 7 x1 y1 x1r y1r phi theta psi
whith:
N Projection number
x1,y1 coordinate of object center in the projection if
the tilt axis were exactly along y
x1r,y1r Coordinate of the object center, if the tilt axis
is rotated in the projection plane.
phi, theta, psi Euler angles of the projection that will result
in a reconstruction of the object oriented along y.
]

.Output doc file for subvolume info: volrotdoc001
[This document file will contain the angles by which the
subvolume needs to be rotated such that the object axis
is along y. These euler angles need to be applied negative
to rotate the volume. Since the rotation is determined only
by 2 points, there is no rotation around the object axis, and
therfore two angles are sufficient, i.e. the third angle is
always 0. If the angles are "alpha, beta, 0" the volume rotation
in spider is with the anlges:
phi=0, theta = -beta, psi=-alpha
]

.enter angular increment of tomogram: 2
[Enter the angular increment used in recording the tilt series.]

.Specify rotation center: 1=first point, 2=second, 3=center: 1
[Enter the point to be used as rotation center. This point
will also afterwards be in the center of the projection and
in the center of the volume.]

.Current tilt-axis angle in the projection: -1.5
[Enter the angle by which the tilt axis is rotated in the 0deg
projection, relative to y. (sign has to be established)]

REMARKS: The preparation of the input document file may require several steps.

1. Often the coordinates are picked from a volume that may have been
interpolated down, in which case this needs to be corrected.

2. The dimensions, first tilt angle, first file number and highest

tilt angle have to be provided manually when creating the input
document file. The following example is primitive and much more
elegant solutions are possible.

3. Typical sequence when the coordinates are picked in imod, using
   etomo:

a) Save the imod coordinates using "model2point -contour" in IMOD

in EMIRA:
b) Use "imodcoordinates2doc" to convert to spider format document file.
   (this also interchanges y and z, see manual).

```
  imodcoordinates2doc
  ../Imod_Coord/emaa_rmlc_tom002.txt   #Input imod coordinates file:
  ../Emacoord/ema_coord002           #Output doc file:
  4                             #number of columns in input (3 or 4):
```

c) Convert the projection stack using "bconvert". (you may use em2em
   but must not use the mirror option.)

d) Use "pickslice" to extract the projections. (the projections coming
   out of bconvert are in a volume, not a stack.)

e) If the coordinates were selected from a reduced volume, then correct
   their values (v.redfac=reduction factor):

```
     v.x1=v.x1*v.redfac
     v.y1=v.y1*v.redfac
     v.z1=v.z1*v.redfac
```

f) The document file that picstickturn needs can then be produced as:

```
v.xydim=2048               #full volume dimensions
v.zdim=600
v.firstangle=-72        #First tilt angle
v.firstpnumber=1        #First image number
v.lasproj=72               #Last projection number

for v.i from 1 to 2 do  #this example is for 2 sticks from the same
                         #tomogram, reads in total 4 lines.
v.line1=(v.i-1)*2+1         #Each stick has 2 coordin. in subsequent lines
v.line2=v.line1+1
readdoc                #read from the coordinate file created with
coodoc001                #imodcoordinates2doc.
v.line1,v.x1,v.y1,v.z1 #read point1
readdoc
coodoc001
v.line2,v.x2,v.y2,v.z2 #read point 2
coodoc001
writedoc                #write the document file for pickstickturn:
windoc001
v.i,v.x1,v.y1,v.z1,v.x2,v.y2,v.z2,v.xydim,v.xydim,v.zdim,\
v.firstangle,v.firstnumber,v.lastangle
enddo

writedocclose
windoc001
readdocclose
coodoc001
end
```

PROGRAMS: tomcoowincalc.f (in fstandalone/applications), pickstickturn (binary)
em_pickstickturn.py in emira/applications

AUTHOR: M.Radermacher

plothisto plot a histogram from a document file to a postscript file

file format: SPIDER document file (input) postscript (output)

PURPOSE: Plot the data in a document file histogram style (does not calculate the histogram).

USAGE: plot histo

.Input file: docfile001
[Enter the name of the document file that contains the data of the histogram. This should be a space separated document file.]

.Postscript file: histopost001
[Enter the name of the postscript file that will have the histogram. Directory will be created if needed.]

.X-column, Y-column: 1,2
[Enter the columns of the document file that should be plotted. Use "0" for the key column. Default 1,2]

.X-label text: x-axis
[Enter the text for the label of the x-axis, default 'X'.]

.Y-label text: x-axis
[Enter the text for the Y-axis label, default 'Y']

.X-range (exampl 1:100): 1:50
[Enter the range of the x-axis labels to be plotted. Default automated choise of range.]

.Y-range (exampl 1:100): 1:400
[Enter the Y-range of the plot. Default automated choice of range.]

.Font size multiplier: 2.
[Enter the relative font size. Default 1]

REMARK: This program may undergo a number of iprovements with usage. This manual may not be complete.

Programs: gnuplot, em_plothisto.py

Author(s): M.Radermacher, others.

Requires: Installation of GNUPLOT

plot lines Plot lines from a document file.

```
inpt file: ASCII, white space separated columns.

REMARK: Since plotting has many different needs, the routine may change. Please
        be not be surprized if these instructions are out-of-date.

USAGE:        plot lines

        .Input file: datadoc001
        [Enter a file with the data you like to plot. This can be any ascii
         file with space separated columns. Since it is not special for
         document files, please make sure you count the column correctly.
         Count also the column that indicates the number of values in a document
         file.]

        .Postscript file: plotpost001
        [Enter the name of the postscript file that should receive the
         plot. You can also add the .ps extension, which may be better
         for later easy printing.]

        .X-label test: absissa
        [Enter the text you would like to write to the X-axis. Default
         is simply X.]

        .Y-label text: ordinate
        [Enter the text you like to use to label the y-axis. Default is
         Y.]

        .X-range (example 1:10): 1:100
        [Enter the range of the x-axis. This range is used to print the
         axis. Should be at least the range of the curve.]

        .Y-range (example 1:10): 0.1:0.8
        [Enter the range of the y-axis. Again this is used for the layout,
         not for the data. If the range is too short, the curve will be
         truncated.]

        .Font size multiplier: 2
        [If you want a larger or smaller font you can specify this here]

        .Number of curves in plot: 2
        [Enter the number of curves you like to plot onto the same sheet.]

For each curve it asks the following questions:
        .X-column, Y-column, Y-error: 1,2,3
        [Enter the data columns. If an error column is specified, the error is
         plotted as error bars. This is at this time too dense, but still
         useful to get an idea.]

        .X-multiplier, Y-multiplier: 1.,1.
        [Specify a multiplier for the data. This is mostly interesting when
        curves of different ranges should be put on the same page.]

Example for plotting an FRC curve from a document file with 50 keys:

        plot lines
        frcdoc001       #Input file:
        nexttry003.ps   #Postscript file:
        Fourier radius  #X-label text:
        FRC     #Y-label text:
        0:0.5   #X-range (exampl 1:100): (here the absolute fourier radius is used)
        0:1.1   #Y-range (exampl 1:100): (the y-axis is left longer by 0.1 for better layout)
        1       #font size multiplier :
        2       #Number of curves in plot (def 1): (plots the FRC and the noise curve)
        3,5     #X-column, Y-column, Y-error: (3 is the abs. radius column, 5 the FRC column)
        1,1     #X-multiplier, Y-multiplier:
        3,6     #X-column, Y-column, Y-error: (3 is the abs. radius column, 6 the noise column)
        1,1     #X-multiplier, Y-multiplier:


Programs: em_plotlines.py, uses guplot
```

Author: M. Radermacher

Requires: Installation of GNUPLOT

plotmlines Plot lines from multiple document files.

```
inpt file: ASCII, white space separated columns.

REMARK: Since plotting has many different needs, the routine may change. Please
        be not be surprized if these instructions are out-of-date.

USAGE:        plot lines

        .Input file: datadoc001
        .Input file: datadoc002
        .Input file: datadoc002
        .Input file: *

        [Enter a file names with the data you like to plot. This can be any
         ascii file with space separated columns. Since it is not special for
         document files, please make sure you counter the column correctly.
         Count also the column that indicates the number of values in a document
         file. Enter all the files from which you want to plot. Indicate end by
         providing "*"]

        .Postscript file: plotpost001
        [Enter the name of the postscript file that should receive the
         plot. You can also add the .ps extension, which may be better
         for later easy printing.]

        .X-label test: absissa
        [Enter the text you would like to write to the X-axis. Default
         is simply X.]

        .Y-label text: ordinate
        [Enter the text you like to use to label the y-axis. Default is
         Y.]

        .X-range (example 1:10): 1:100
        [Enter the range of the x-axis. This range is used to print the
         axis. Should be at least the range of the curve. NOTE: THE
         FORMAT FOR THE RANGE IS A:B, ITS SEPARATED BY ":". The most common
         reason for an error is a mistake in this input format.]

        .Y-range (example 1:10): 0.1:0.8
        [Enter the range of the y-axis. Again this is used for the layout,
         not for the data. If the range is too short, the curve will be
         truncated.NOTE: THE FORMAT FOR THE RANGE IS A:B, ITS SEPARATED BY ":". The most common
         reason for an error is a mistake in this input format.]

        .Font size multiplier: 2
        [If you want a larger or smaller font you can specify this here]

        .Number of curves in plot: 2
        [Enter the number of curves you like to plot onto the same sheet.
         The number of curves are assumed to be the same for all files.]

For each curve it asks the following questions:
        .X-column, Y-column, Y-error: 0,1,2
        [Enter the data columns. Column 0 is the same as the key of the
         document file. If an error column is specified, the error is plotted
         as error bars. This is at this time too dense, but still useful to get
         an idea. The colums are assumed to be the same in all files.]

        .X-multiplier, Y-multiplier: 1.,1.
        [Specify a multiplier for the data. This is mostly interesting when
        curves of different ranges should be put on the same page.]

Example for plotting an FRC curve from a document file with 50 keys:

        plot lines
        frcdoc001       #Input file:
        frcdoc002       #Input file:
        frcdoc003       #Input file:
        *               #Input file:
        nexttry003.ps   #Postscript file:
        Fourier radius  #X-label text:
```

```
FRC     #Y-label text:
0:0.5   #X-range (exampl 1:100): (here the absolute fourier radius is used)
0:1.1   #Y-range (exampl 1:100): (the y-axis is left longer by 0.1 for better layout)
1       #font size multiplier :
2       #Number of curves in plot (def 1): (plots the FRC and the noise curve)
3,5     #X-column, Y-column, Y-error: (3 is the abs. radius column, 5 the FRC column)
1,1     #X-multiplier, Y-multiplier:
3,6     #X-column, Y-column, Y-error: (3 is the abs. radius column, 6 the noise column)
1,1     #X-multiplier, Y-multiplier:
```

Programs: em_plotmlines.py, uses guplot

Author: M. Radermacher

Requires: Installation of GNUPLOT

plotscatter2d Create a 2D scatter plot

file format: ASCII space separated columns.

USAGE: plot scatter 2d

```
        .Input file: datacolumns001
        [Enter name of input file. This can be a document file or any other
        file containing data columns. Since it is not special for document
        files, column 1 in a traditional document file will need to be addressed
        as column 3 in this program. The columns can be white space separated
        of by a character that can be specified, e.g. comma.]

        .Column separator: ,
        [If anything else but whitespace separates the column, specify the
        separator here. Otherwise hit return.]

        .Output file format (PNG or PS): ps

        .Postscript file: Scatterplot001.ps
        [Enter postscript file name. It is recommended to specify the extension
         for easier usage later.]

        .X-label text: abscissa
        .Y-label text: second coordinate
        [Enter the labels that will be printed along the axis. Defaults are
         X and Y]

        .X-range (exampl 1:100): -50:60
        .Y-range (exampl 1:100): -10:45
        [Specify the range of the values for scaling of the axis. If no
         values are provided the plot will be autoscaled. Watch for the
         correct format, from to to are separated by a ":" ]

        .Label axes y/n: y
        [Specify if labels (numbers) should be written on the axes.]

        .Grid y/n: y
        [Specify if the plot should show a grid.]

        .Linewidth (def 1): 1
        [Specify line width on the plot.]

        .font size multiplier : 2
        [Specify a number >1 if you need larger fonts.]

        .X-column, Y-column, z-column,labelcolumn: 1,2,3,1
        [Enter which columns should be used for plotting. Answers are in the
         order: x-coordinate, y-coordinate, followed by the
         column that should be used to label the points.]

        .Multipliers for X,Y,Z: 1,1
        [If you like to scale the input values, specify this here.]
```

Programs: em_plotscatter3d.py, gnuplot

Author: M.Radermacher (of the python code)

Requires: Installation of GNUPLOT

plotscatter3d **Create a 3D scatter plot**

file format: ASCII space separated columns.

USAGE: plot scatter 3d

```
        .Input file: datacolumns001
        [Enter name of input file. This can be a document file or any other
        file containing data columns. Since it is not special for document
        files, column 1 in a traditional document file will need to be addressed
        as column 3 in this program. The columns can be white space separated
        of by a character that can be specified, e.g. comma.]

        .Column separator: ,
        [If anything else but whitespace separates the column, specify the
        separator here. Otherwise hit return.]

        .Postscript file: Scatterplot001.ps
        [Enter postscript file name. It is recommended to specify the extension
         for easier usage later.]

        .X-label text: abscissa
        .Y-label text: second coordinate
        .Z-label text: z-direction
        [Enter the labels that will be printed along the axis. Defaults are
         X,Y and Z.]

        .X-range (exampl 1:100): -50:60
        .Y-range (exampl 1:100): -10:45
        .Z-range (exampl 1:100): -0.1,0.5
        [Specify the range of the values for scaling of the axis. If no
         values are provided the plot will be autoscaled.]

        .font size multiplier : 10
        [Specify if you need larger fonts.]

        .X-column, Y-column, z-column,labelcolumn: 1,2,3,1
        [Enter which columns should be used for plotting. Answers are in the
         order: x-coordinate, y-coordinate and z-coordinate followed by the
         column that should be used to label the points.]

        .Multipliers for X,Y,Z: 1,1,1
        [If you like to scale the input values, specify this here.]
```

Programs: em_plotscatter3d.py, gnuplot

Author: M.Radermacher (of the python code)

Requires: Installation of GNUPLOT

PPCAEM PROBABILISTIC PRINCIPLE COMPONENT ANALYSIS WITH EXP. MAX.

file format: SPIDER

SEE ALSO: PLOTSCATTER3D

REMARK: The program has been mainly used for Fourier-Radon transforms of volumes. Prerequisite are aligned volumes. If they are in real space, first Radon transform them, the Fourier transform. These volumes will not have a missing data indes, Therefore the fraction of missing data should be provided so that the program finds the missing data by analyzing the sigmas of the radial lines.
It is better,if the Radon transforms of the volumes are calculated directly from the projections in the proper orientation. Then the missing data index is present.The latter is more precise. The projection angles for a rotated volume can be calculated with "combineuler" (vo ceul in SPIDER).
If a carthesian Fourier is used, an index file for the missing data must be provided.The format of this file is not clear at the time I am writing this manual, Please consult the c-code of the program.

USAGE: PPCAEM

.First file name: volume001
[Enter the name of the first volume containing densities.
CAUTION: this program replaces all numbers at the end of a file name with a 3-DIGIT number. Until the c-program is changed, the volumes names must have 3 digits (not more, or less).]

.Image gaps allowed) first,last image number: 1,50
[Enter the first and l;ast number of the volume file series.]

File type options:
. Real
. Radon
. Fourier
. polar_Fourier
(first 2 letters sufficient, case insensitive)

.Input file type: polar
[Enter the file type of the input data. Polar Fourier transforms, obtained from 3D Radon transforms have been used most and should be the most reliabe. Note that the polar transforms must have the information on the line variance and the averaging counter in the last for numbers. (M.R.'s Radon/Radon Fourier format).

.Mask file or *: mask001
[Enter the name or * if not mask is applied.]

.Radius, abs. Fourier or real space pixels: 0.08
[Enter the radius for a spherical mask in aither Fourier or real space. This value is ignored if a mask file was provided]

.Fraction of missing data: 0.33
[Enter the fraction of data that is missing. This value is required if the information is not contained in the input data. Then the missing radial lines are identified based on the statistics (sigma) of each radial line.]

The following file is read if it exists,
otherwise it is created.
.
.File with file list: filelist001
[Enter name fo file where the list of files is kept.]

.File name for data matrix: ppcaemmatrix001
[Enter a file name for storing the data matrix.]

.File name for data index: indexfile001
[Enter a file name where the indices of existing/non-existing data are kept.]

```
     ALL THE BOVE IS USED IN THE PREPARATION STEP.
     nOW FOR THE ACTUAL ppcaeM:

               .Number of factors: 8
          [Enter the number of factors to be calculated in PPCAEM.]

          .Number of iterations: 1000
          [Enter the number of iterations in the expectation maximization.
           default: 1000]

          .Conversion criterion: 0.00001
          [Threshold to stop interations. Default 0.00001]

          .Log file for PPCAEM: ppcaemlog001
          [Enter a name for the log file of ppcaem.]

          .Coordinate file: Outcoordinates001
          [Enter a name for the coordinate output file.]

          .Name for coordinate file with file numbers: IDcoordinates001
          [This will contain the coordinates identified with the
           volume file number.]

          .Eigen vector file: eig001
          [Will receive the eigenvectors]

          .Eigenvalue file: eigval001
          [Enter the name of the file to receive the list of eigenvalues.]

          .Estimates file: estimates001
          [Contains estimated missing data.]

          .Example of output file series: Reconstvol001
          [Recreates the volumes with restored missing data.]

          .Example for Eigenvector file series: eigenvol001
          [Creates the volumes that correspond to each Eigenvector.]

          .Enter step to run (0=all): 0
          [Enter which step of the process to run. This is useful if
          one step went wrong, enabling the use of previous steps.
          The options are:
          0 = run all
          1 = run only the preparation step (prepca)
          2 = run only ppcaem
          3 = run only postpca (create eigenvolumes, reconstituted volumes etc)
          ]

Notes:    This program runs prepca, ppcaem and postpca. For more details
          see the PPCAEM directory with Lingbo Yu's instructions for
          command line use.

Programs: prepca, postpca, ppcaem

Author(s): M. Radermacher: for .py, Lingbo Yu for the PPCAEM programs
           written in C.

See:            L.Yu et al. J. Struct. Biol. 171, 2010, 18-30.

Below are examples from Lingbo's original:

#The answers for running ppcaem_batch
#!!!!!!!!!!!DO NOT CHANGE THE VARIABLE NAMES!!!!!!!!!!!
##-------------------START of user inputs-------------------
# From which step you want to run this batch
STEP=1
#
EXTENSION="spi"
## The input files must be polar Fourier volumes
EXAMPLE="handalignedradft001"
## The program allows gaps in your file name sequence,
##  so please provide the first and the last number
```

```
FIRST=1
LAST=80
RADIUS=0.08
## The percentage of missing data: 0 if no missing data
##  -1 if the information is contained in the original
##  polar Fourier volumes
PERCENT=0.3
NOFACTORS=10
MAXITER=600
CONVERGE=1e-6
## Outputs
OUTPUT_VOLUMES="handalignedradftest001"
OUTPUT_COORDINATE="ppca_coord"
OUTPUT_EIGENVALUES="ppca_eigval"
OUTPUT_EIGENVOLUMES="eigenvector001"
## A file (output) contains a list of the volumes
##  contributing to the columns of the data matrix
OUTPUT_FILELOG="handaligned_filelog"
##-------------------END of user inputs-------------------

AND HERE IS AN EXAMPLE BATCH SCRIPT TO RUN THE GENERIC (NOT THROUGH EMIRA)
PPCAEM:

#!/bin/sh
getopt_in()
{
        if [ "$1" ]
        then
                if [ ${1:0:1} = '-' ] && [ ${1:1} = 'i' ]
                then
                        IFLAG=2
                else
                        IFLAG=1
                        ANSWER=$1
                fi
        else
                IFLAG=0
        fi
}
echo "Hello! This is a batch file to run PPCA-EM"
echo "Two ways to run: "
echo "  1, To run interactively, use the argument \"-i\""
echo "  2, To run with anwers in a file, use the file name as the argument"
echo "     (an example will be generated when first run interactively)"
echo "Four steps to run: "
echo "  1, prepca -- prepare two matrices, data and index"
echo "  2, ppcaem -- run ppca-em"
echo "  3, postpca -- save the output matrices back to volumes"
echo "  4, gnuplot -- generate pairwise scatterplots"
echo "Notes: The inputs are polar Fourier transforms of your volumes,"
echo "which can be acquired by running \"rm 3dn\" and \"rm ftn\" in SPIDER."
echo "Authers: Lingbo Yu (lyu@uvm.edu)"
echo
getopt_in $*
if [ ${IFLAG} -eq 1 ]
then
#echo $ANSWER
source $ANSWER
echo "The inputs are:"
echo "STEP=$STEP"
echo "EXTENSION=$EXTENSION"
echo "EXAMPLE=$EXAMPLE"
echo "FIRST=$FIRST"
echo "LAST=$LAST"
echo "RADIUS=$RADIUS"
echo "NOFACTORS=$NOFACTORS"
echo "PERCENT=$PERCENT"
echo "MAXITER=$MAXITER"
echo "CONVERGE=$CONVERGE"
echo "OUTPUT_VOLUMES=$OUTPUT_VOLUMES"
echo "OUTPUT_COORDINATE=$OUTPUT_COORDINATE"
echo "OUTPUT_EIGENVALUES=$OUTPUT_EIGENVALUES"
echo "OUTPUT_EIGENVOLUMES=$OUTPUT_EIGENVOLUMES"
```

```
echo "OUTPUT_FILELOG=$OUTPUT_FILELOG"
elif [ ${IFLAG} -eq 2 ]
then
#if run iteractively
echo -n -e "From which step you want to run this batch:\n"
read STEP
echo -n -e "The file extension:\n"
read EXTENSION
echo -n -e "An example filename of the polar Fourier volumes:\n"
read EXAMPLE
echo -n -e "The number of your first volume:\n"
read FIRST
echo -n -e "The number of your last volume:\n"
read LAST
echo -n -e "The Fourier radius:\n"
read RADIUS
echo -n -e "The number of eigenvectors:\n"
read NOFACTORS
echo -n -e "The maximum number of iterations:\n"
read MAXITER
echo -n -e "The criteria for convergence, 1e-6 for example:\n"
read CONVERGE
echo -n -e "The percentage of missing data, 0.3 for example, -1 if the informatio is contained in the polar I
read PERCENT
echo -n -e "An example filename of the output polar Fourier volumes, with the missing data estimated:\n"
read OUTPUT_VOLUMES
echo -n -e "The filename for the output coordinates:\n"
read OUTPUT_COORDINATE
echo -n -e "The filename for the output eigenvalues:\n"
read OUTPUT_EIGENVALUES
echo -n -e "An example filename of the output eigenvolumes:\n"
read OUTPUT_EIGENVOLUMES
echo -n -e "The filename for an output file with the numbers of available volumes:\n"
read OUTPUT_FILELOG
#save the inputs
ANSWER="#The inputs of run_ppcaem:"
ANSWER=${ANSWER}"\n# while changing the answers, don't change the VARIABLE names!!"
ANSWER=${ANSWER}"\n#"
ANSWER=${ANSWER}"\n# From which step you want to run this batch\nSTEP="${STEP}
ANSWER=${ANSWER}"\n# The file extension\nEXTENSION="${EXTENSION}
ANSWER=${ANSWER}"\n# An example filename of the polar Fourier volumes\nEXAMPLE="${EXAMPLE}
ANSWER=${ANSWER}"\n# The number of your first volume\nFIRST="${FIRST}
ANSWER=${ANSWER}"\n# The number of your last volume\nLAST="${LAST}
ANSWER=${ANSWER}"\n# The Fourier radius\nRADIUS="${RADIUS}
ANSWER=${ANSWER}"\n# The number of eigenvectors\nNOFACTORS="${NOFACTORS}
ANSWER=${ANSWER}"\n#The percentage of missing data\nPERCENT="${PERCENT}
ANSWER=${ANSWER}"\n# The maximum number of iterations\nMAXITER="${MAXITER}
ANSWER=${ANSWER}"\n# The criteria for convergence\nCONVERGE="${CONVERGE}
ANSWER=${ANSWER}"\n# An example filename of the output polar Fourier volumes\nOUTPUT_VOLUMES="${OUTPUT_VOLU
ANSWER=${ANSWER}"\n# The filename for the output coordinates\nOUTPUT_COORDINATE="${OUTPUT_COORDINATE}
ANSWER=${ANSWER}"\n# The filename for the output eigenvalues\nOUTPUT_EIGENVALUES="${OUTPUT_EIGENVALUES}
ANSWER=${ANSWER}"\n# An example filename of the output eigenvolumes\nOUTPUT_EIGENVOLUMES="${OUTPUT_EIGENVOL
ANSWER=${ANSWER}"\n# The filename for an output file with the numbers of available volumes\nOUTPUT_FILELOG="
echo -e ${ANSWER} > _run_ppcaem_inputs_
fi
##################################################
#starting the real process
if [ ${IFLAG} -gt 0 ]
then
echo
date
echo "Let's start working"
#step 1: prepca
if [ ${STEP} -eq 1 ]; then
        echo ${STEP}", prepare a data matrix and an index matrix."
        SIGN=`echo ${PERCENT} |grep "-" |wc -l`
        if [ ${SIGN} -eq 0 ]; then
                PPCAEM_CMD="prepca -x ${EXTENSION} -e ${EXAMPLE} -f ${FIRST} -l ${LAST} -t polar_Fourier -r
                PPCAEM_CMD=${PPCAEM_CMD}"-d _ppca_data_ -i _ppca_index_ -g ${OUTPUT_FILELOG}"
        else
                PPCAEM_CMD="prepca -x ${EXTENSION} -e ${EXAMPLE} -f ${FIRST} -l ${LAST} -t polar_Fourier -r
                PPCAEM_CMD=${PPCAEM_CMD}"-d _ppca_data_ -i _ppca_index_ -g ${OUTPUT_FILELOG}"
        fi
```

```
        echo "..."${PPCAEM_CMD}
        ${PPCAEM_CMD}
        STEP=$[${STEP}+1]
fi
#step 2: ppcaem
if [ ${STEP} -eq 2 ]; then
        echo ${STEP}", running PPCA-EM. (This may take a while, have a cup of tea!)"
        PPCAEM_CMD="ppcaem -x ${EXTENSION} -d _ppca_data_ -i _ppca_index_ -f ${NOFACTORS} -m ${MAXITER} -g $
        PPCAEM_CMD=${PPCAEM_CMD}"${OUTPUT_COORDINATE} -v _ppca_eigvec_ -e ${OUTPUT_EIGENVALUES} -t _ppca_es
        echo "..."${PPCAEM_CMD}
        ${PPCAEM_CMD}
        STEP=$[${STEP}+1]
fi
#step 3: postpca
if [ ${STEP} -eq 3 ]; then
        echo ${STEP}", saving the results"
        PPCAEM_CMD="postpca -x ${EXTENSION} -e ${EXAMPLE} -d _ppca_est_ -v _ppca_eigvec_ -t polar_Fourier -r
        PPCAEM_CMD=${PPCAEM_CMD}"-o ${OUTPUT_VOLUMES} -f ${FIRST} -l ${LAST} -g ${OUTPUT_EIGENVOLUMES}"
        echo "..."${PPCAEM_CMD}
        ${PPCAEM_CMD}
        STEP=$[${STEP}+1]
fi
#step 4: gnuplot
if [ ${STEP} -eq 4 ]; then
        echo ${STEP}", generating the scatter plots. (The gnuplot scripts are temporarily stored in _gnuplot_
        #plot scatter plots
        PLOTSETTING="set datafile separator \",\"\\n"
        PLOTSETTING=${PLOTSETTING}"set datafile commentschars \";\"\\n"
        PLOTSETTING=${PLOTSETTING}"set key off\\n"
        PLOTSETTING=${PLOTSETTING}"unset mouse\\n"
        PLOTSETTING=${PLOTSETTING}"set term postscript enhanced color solid font \"Times,8\"\\n"
        for i in `seq 1 $[${NOFACTORS}-1]`;
        do
                PLOTSETTING=${PLOTSETTING}"set xlabel \"Principal Component "${i}" (Real)\"\\n"
                i1=$[$[${i}*2]+1]
#               i2=$[${i1}+1]
                for j in `seq $[${i}+1] ${NOFACTORS}`;
                do
                        j1=$[$[${j}*2]+1]
#                       j2=$[${j1}+1]
                        PLOTSETTING=${PLOTSETTING}"set output \"ScatterPlot_${i}_${j}.ps\"\\n"
                        #echo "...plotting principal component ${i} v.s. ${j}"
                        PLOTSETTING=${PLOTSETTING}"set ylabel \"Principal Component ${j} (Real)\"\\n"
                        PLOTSETTING=${PLOTSETTING}"plot \" _gnuplot_script_
        PPCAEM_CMD="gnuplot _gnuplot_script_"
        echo "..."${PPCAEM_CMD}
        ${PPCAEM_CMD}
fi
else
echo "WRONG INPUT! PLEASE READ THE INSTRUCTION CAREFULLY"
fi
```

predictresolution predict resolution for conical or single axis 3D reconstr.

file format: SPIDER document file

USAGE: predict resolution

```
          .Output doc file: resolpredict001
          [Enter the name of the output document file that will contain
           the resolution depending on the tilt increment and object
           diameter. ]

          .First, last tilt increment, stepsize: 0.5,5,0.5
          [Enter the smalles and largest tiltincrement and the steps. This
           will calculate a curve of resolution versus increment.]

          .Increment for which values should be returned: 3
          [Enter a specific increment value for which you would like the value
           returned in a variable.]

          .Object diameter: 110
          [Specify the object diameter, assumed to be spherical.]

          .S (single axis) or C (conical):s
          [Specify if you would like the resolution value for single axis
           or conical tilting.]

          .Enter variable to receive output, (def=v.predictres) :v.predict
          [Enter the name of the variable to contain the specific answer. This
           variable will contain: "v.predict.resol" resolution in the same
           units are the specified diameter. "v.predict.reciprocal" the reci-
           procal value of the resolution.]
```

Remark: For the calculation of the single axis tilt resolution Crowther's formular is used.
For the calculation of the conical resolution Radermacher's formula (for even number of projections only) is used
(see Radermacher, Hoppe 1978,1980 and Radermacher M., Thesis 1980)

programs: restomocurve.f
Author: M.Radermacher

project **project a 3D volume onto 2D projection**

file format: SPIDER

USAGE: project

.Input file: volume001
[Enter the name of the volume you want to project.]

.Output file: projection001
[Enter the file name of the projection]

.X, Y-dimensions of projection: 128,128
[specify the dimensions of the projection.]

.phi, theta, psi: 10.,12.5,31.
[Enter the Euler angles of the projection. Phi around Z,
theta around Y', psi around Z'' = in-plane rotation.]

Programs: em_project.py, pj3_n.f

Author(s): M.Radermacher (.py and pj3.f, pj3_n.f standalone version),
P.Penczek (pj3-rewrite: pj3_n.f subroutine)

qvol vector quantization of volume (SITUS program)

file format: MRC or SITUS

USAGE: qvol

```
        .Input file name: volume001.mrc
        [Enter the input volume. Format either MRC or SITUS]

        .Vector file: vectors001.pdb
        [Enter the file for the vector coordinates. It is written
         in PDB format and can directly be used for example by Chimera.]

        .PSF file: psf001
        [Enter name for the psf file. (meaning see below).]

        .Connectivity file: connections001.con
        [Enter file name for vector connectivities.]

        .Enter density options (1,2, or 3): 2
        [Option 1 does not do anything, option 2 writes density statistics
         to the Results file, option 3 calculates the volume histogram and
         writes it to the Results file. (In SITUS this information is printed
         to the screen. We will improve the wrapping so that in the future
         also EMIRA can send the output to the terminal, if desired)]

        .Cutoff density value: 0.1
        [Enter the density value that determines the boundaries of the
         molecule.]

        .Number of codebook vectors: 15
        [Enter the number of codebookvectors that are to be determined.]
```

Remarks: The program requires SITUS in the program path.
The implementation at this ime is "proof of principle", therfore not all options (e.g. a start vector file) are provided. This will change in the future).

COPY OF SITUS INSTRUCTIONS (http://situs.biomachina.org/fguide.html):

qvol - Vector Quantization of Volumetric Map

Purpose:

Specialized tool to perform a vector quantization of low-resolution, single molecule data. qvol supports the correlation-coefficient based docking with qdock, and flexible docking with qplasty. In the absence of existing vector positions, qvol carries out a global search using the TRN algorithm. If start vectors are already known, the LBG local search algorithm is used instead of TRN, or connectivities can be learned. LBG allows to add distance constraints to the vector refinement that are useful for flexible docking.

With TRN, a small number of calculations (8 by default) are repeated with different random number seeds. The averaged codebook vectors and their statistical variability are then written to the output file. With LBG, no statistical clustering is performed. In this case it is important to specify reliable initial positions from a prior qvol run.

Usage:

In a practical application of qvol, one should extract from the volumetric data a region of interest corresponding to a single molecule using e.g. voledit. Next, the user must determine a suitable number of codebook vectors. Only densities above a user-defined threshold value are considered by qvol to eliminate background noise in the low-resolution data. Depending on the noise, this threshold value should be at 50-80% of the level that is typically considered the "molecular surface" of the biopolymer in the low-resolution data.
New vector positions are calculated automatically with the TRN method if no start vectors are specified. Subsequently, these vector positions can be refined in a second qvol run with the LBG method. Also, any distance constraints can be read from a file or entered at the command prompt at this time.

The input map must be in Situs or CCP4/MRC format (auto detect). To convert
other maps to either of these formats use the map2map utility.

Usage (at shell prompt):
./qvol file1 [file2] file3

file1: inputfile, density map
file2: inputfile, start vectors, PDB format (optional)
file3: outputfile, PDB format

Interactive input at program prompt (also suitable for automation):

    Choice of utilities to inspect the density distribution (e.g. voxel
     histogram).
    Threshold (cutoff) density value.
    Number of codebook vectors.
    (If file2 is specified): Choice of entering distance constraints
    manually or from a file. There are two constraint file options.
    Constraint file entries generated e.g. with qpdb are triples of
    free-format values in the order   ,
    where the indices correspond to the order of vectors in file2, counting
    from 1. It is also possible to read the connectivities from a PSF file
    in which case the missing distances are computed from file2.
    Choice of computing the vector connectivities (neighborhood relationships)
    with the Competitive Hebb Rule and writing them to a file.
    You can automate this interactive program by "overloading" the
    standard input (if you put expected values in a script, see
    run_tutorial.bash script in the tutorials).

Output:

    (Program level:) Statistical analysis of the vectors and their radius
    of gyration, i.e. the radial rms deviation from the vector center of mass.

    (Shell level): Codebook vectors in a PDB-formatted output file. The
    vector rms variabilities, representing the precision of the codebook
    vectors, are written to the occupancy fields of the PDB-style atom entries. (Optional) Vector connectivi

Notes:

    Vector connectivities in PSF format can be visualized and edited as
    bond connections (together with the atom-style PDB entries of file2
    and file3) using the molecular graphics program VMD. Simply overload
    the PSF file into the PDB file in the VMD 'Molecule' menu. Then under
    the 'Mouse' menu select 'Add/Remove Bonds'. The edited connectivity
    can then be saved later into a PSF file from the VMD command console
    (assuming your molecule is 'top'):

    set sel [atomselect top all]
    $sel writepsf my.psf

    If there are cluster size deviations from the expected value (default: 8)
    when using the TRN algorithm, refine the found vector positions by
    passing them to qvol as input file of a second, LBG run.

    Distance constraints do not determine the chirality (handedness) of
    vector connections. If you encounter mirror images or otherwise
    flipped connections after running qvol compared to connections
    determined with qpdb, you need to experiment with the indexing
    of your constraints. The LBG method combined with the SHAKE constraint
    algorithm is relatively insensitive to the position of start vectors.

Requires: Installation of SITUS

rad2d Calculate a 2D Redon transform

image format: SPIDER

USAGE: rad2d

```
        .Input file: image001
        [Enter name of 2D image file.]

        .Output file: radon2d001
        [enter name for the 2D raodn transform output file]

        .P-dimension of Radon transform: 100
        [Enter the radial dimension of the Radon transform.]

        .Mask raduis, angular increment: 55.,2.5
        [Enter the radius of the (smooth) mask that is applied before
         calculating the Radon transform. Unfortunately a circular mask is
         necessary to avoid the Radon transform from being influenced by the
         shape of the image area. The outside of the mask is "0". Therfore
         it is important to set the image average (i.e. the average value of
         the area in the image where there is no object) to 0.0. Otherwise a
         step may be created which prevents the use of the Radon transform for
         alignments.
         Enter the angular increment you want to use. The Radon transform
         will be calulated wintin an angular range from -90 to +90-increment.
         Note: the rest of the transform is mirror related to this half
         transform and therfore does not need to be calculated.]

        .Center offset in x,y: 0,0
        [ This option allows to move the rotation center that is used in
          calculating the Radon transform. Normally it is left at 0,0]

         .Threshold: Average threshold, Threshold, Lift, None (a/t/l/n): n
         [Specify if the image should be thresholded in any way during the
          calculation of the Radon transform. Options are: Use the image
          average as threshold (everything below is set to threshold), specify
          two threshold values (see below), lift all image values so that all
          are positive, or None. Thes options are mainly provided
          for experimental reasons and not used in every day processing.]

         If the option T (=threshold) was used:

                .upper, lower threshold: 1.,0.
                [Enter upper and lower threshold. Values outside will be set
                 to these values.
```

Programs: em-rad2d.py, rad2dcalln.f, radon2en.f

Author: M. Radermacher

rad2d3dali align a volume to a reference projections

image format: SPIDER

USAGE: rad2d3dali

```
        .3D Fourier Radon transform: rad3df001
        [Input the name of the 3D transform that is being aligned.]

        .Output F.Radon sections (or *):*    (or sect000001)
        [If a file name is given all section extracted from the 3D
         Fourier Radon transform are written to files. (for a search
         range -180-180,10 in phi, -90,90,10 in theta and -180,180,10
         in psi these will be over 23000 files.)

         .Cross-correlation peak file: peak0001
        [Enter the name to keep the angular cross-correlation function.]

        .Shift radius file (or *): shiftrad0001
        [This file will contain the shift length for every orientation.
         (The cross-correlation file conatains the coefficient where
          this shift value results in a maximum correlation.)]

        .Shift direction file (or *): shiftang001
        [Enter the name of the file that contains the angle (dhift direction)
         for each value in the shift radius file.]

        .Search phi: from, to, increment: -180,180,10
        [Specify the seach range and incrment in phi, the rotation
         around the volume's z-axis.]

        .Search theta: from, to, increment: -90,90,10
        [Specify the search range and increment for theta, the negative
         rotation around the y'-axis (after phi rotation).]

        .Search psi: from, to, increment: -180,180,10
        [Specify search range and increment for psi, the rotation around
         the z''-axis (after phi and theta rotation).]

        .Shift range (or 0): 5
        [Enter the shift search range. This is the length of the shift and
         determines how much a projection may move relative to the volume.]

 If a shift range was specified:
             .Use shift only in one direction X,Y, N=no: N
              [If X or Y is specified then the shift direction will be
               restricted to either X or Y within the projection.]

        .Skip missing volume data (Y/N): Y
        [Enter 'Y' if your volume has missing data. The cross-correlation
         will then correctly calculate the correlation only from data that
         exist.]

        .Min. % required, threshold for sigma: 30.0,0.0000001
        [Enter the minumum percentage of data that must exist in the cross-
         correltion. If the overlap is less the correlation will be 0. There
         are two ways to recognize missing data. The best is the averaging index
         in the 3D Radon transform, that is created in the reconstruction
         process. If this index is present, the sigma threshold will be
         ignored. If this index is not present, then the threshold determines
         the value below which a line is considered missing data. To determine
         the best value use the command "radsigstat".

Programs: em_rad2d3dali.py, rad2d3dalign.f and subroutines also used in other
          programs.

Author: M. Radermacher
```

rad2dinv Invert 2D Radon transform

```
image format: SPIDER

USAGE:  rad2dinv

        .Input file: rad2d001
        [Enter name of the input 2D Radon transform to be inverted. May be
         Fourier or real]

        .Input file is (R)eal or (F)ourier?: R
        [Enter if input radon transform is in real space of Fourier space.]

        .Output file: radrecon001
        [enter name for the output 2D real space reconstruction.]

        .Dimensions (x,y): 128,128
        [Enter the dimension of the real space image.]

        weighting function options: s:(sqrt(r*)), r:(r*), c:(const)
        .Type of weighting function: r
        [Enter which weighting scheme should be used in the different steps.
         If the 2D Radon transform is in real space, and the projections have
         not been prefiltered with r* or sqrt(r*), then use filter r.
         If the Radon transforms had been prefiltered with
         sqrt(r*), a step advantageous in projection alignment, then the
         inversion only needs a second sqrt(r*) to achieve the
         r* weighting.
         The option 'c' is provided if one want to use a simple
         back-projection without weighting for the inversion.

        .Doc file for weighting fct or *: *
        [Enter a file name if you like to look at the weighting function
         (it is a one-dimensional function). Enter * if you do not want it
         saved.

Programs: em-rad2dinv.py, callrminvn.f, rminvn.f, fouradn.f, fouradinvn.f
          rmiwf.f

Author: M. Radermacher
```

rad3d Calculate a 3D Radon transform

```
image format: SPIDER

USAGE:  rad3d

        .Input file: vol001
        [Enter name of 3D volume file.]

        .Output file: radon3d001
        [enter name for the 3D Radon transform output file]

        .p-dimension of Radon Transform: 128
        [Enter the radial (p) dimension of the output Radon transform. Does not need to
         be the same as the image dimension.]

        .Mask radius Angular Inc. Theta, Phi: 90.,1.5,1.5
        [Enter the radius of the spherical mask used in the calculation of the Radon transform
         and the angular increments of Theta (angle around y-axis) and Phi (in-plane angle
         of each projections with angle theta. To calculate a 3D Radon transform, first a
         single axis tilt series around y (Theta) is calculated, and then in each of the
         projections of this tilt series a 2D Radon transform is calculated. The 2D transforms
         are stacked with Theta to form the 3D Radon transform.]

         .Write intermediate projections: N
         [The intermediate projections can be written out before they are Radon transformed.
         Enter Y (yes) or N (no)]

         If the answer was "Y" (yes) then:
                .1rst projection file name: pro001
               [Enter the first name of the output series. The series will be numbered from 1 to
                the number projections calculated (180 for 1 deg theta increment).]

        .3D threshold options: Threshold,Lift,None (T/L/N): N
        [Specify if you want to apply a threshold to the input 3D. This is mostly interesting
         for model calculations. Option T will ask for lower and upper threshold, Option L
         adds a constant to make the complete volume be positive, N= no threshold (default).]

         If the anser was T:
                .Upper, Lower thresholds: 0.,1.
               [Enter upper and lower thresholds. Defaults are 0. and 1.]

        .2D average threshold (Y/N): N
        [Enter of you would like to threshold the intermediate projections at their average.
         Default N. Also an option implrmrnted for model calculations.]

Programs: em-rad3d.py, radon3dn.f, radon2en.f

Author: M. Radermacher
```

rad3dcopycounter copy counter of 3D Radon transfomr

file format: SPIDER, Radon transform with counter space

USAGE: rad3d copy counter

.Input file: rad3d001
[Enter input Radon transform. The counter found at the end of the lines will be copied to the counter position in the output 3D Radon transform.]

.Output file: otherrad3d001
[This 3D Radon transform must exist. The counter from the first file is copied to the counter positions in this output file.]

remark: This operation is usefull if the counter in a 3D Radon transform got lost and needs to be restored. The output file must be in the same orientation as the input file, otherwise the counters do not make sense. The counters are used among other things to indicate present and missing data.

Programs: rmcountcopy.f

Author(s): M.Radermacher.

rad3dinv Invert 3D Radon transform

```
image format: SPIDER

USAGE:  rad3dinv

        .Input file: rad3d001
        [Enter name of the input 3D Radon transform to be inverted. May be
         Fourier or real]

        .Input file is (R)eal or (F)ourier?: R
        [Enter if input radon transform is in real space of Fourier space.]

        .Output file: radrecon001
        [enter name for the output 3D real space reconstruction.]

        .Intermediate file (use _filename to force it): interfile001
        [Enter the name for an intermediate file, where the 3D is stored
         after the first step of inversion. This file is only needed if
         memory is insufficient. Preceed the name with an underline, if you
         would like to create this file independent of memory needs. This
         can be useful for analyzing the behavious of the inversion (or for
         curiousity).

        .Dimensions (x,y,z): 128,128,128
        [Enter the dimension of the output reconstruction.

        weighting function options:
        1 (const), 2 (sqrt(r*)), 3 (r*)
        .Weighting functions for step1,step2: 3,3
        [Enter which weighting scheme should be used in the different steps.
         If the 3D Radon transform is in real space, and the projections have
         not been prefiltered with r* or sqrt(r*), then use filter 3 in both
         steps (r* weighting). If the projections had been prefiltered with
         sqrt(r*), a step advantageous in projection alignment, then the
         first part of inversion only needs a second sqrt(r*) to achieve the
         r* weighting. In this case filter 2 for step 1. Step 2 ususaly uses
         3 (r*-weighting).
         You can use 1, if a simple backprojection should be used in the
         inversion.

        .Doc file for weighting fct or *: *
        [Enter a file name if you like to look at the weighting function
         (it is a one-dimensional function). Enter * if you do not want it
         saved.

Programs: em-rad3dinv.py, rad3dinvers (binary of call3drminvn.f)

Author: M. Radermacher
```

radalign2d 2D image alignment with Fourier-Radon transforms

image format: SPIDER

USAGE: radalign2d

.Input reference Radon Fourier file: refradf001
[Enter the name of the reference Fourier-Radon transform. This reference should be already appropriately filtered. Minimum filter required in "radfour" is filter 8 (=sqrt(r*)). This high-pass filter in applied to both, the single images adn the reference. With this filter the crosscorelation function has the appropriate weights and sharpness. (alternatives of filtering one with r* instead should be tried first.]

.Input Radon Fourier file example: ../rad2df/imgradf001
[Enter the first file in the series. It also needs to be a Radon Fourier transform, highpass filtere with sqrt(r*) (see radfour).]

.Output document file: radaldoc001
[Enter the name of the output document file to contain the translations and rotations.]

.Output 3D CCF or *: ../ccf/peakfile001
[Enter the name for the 3D crosscorrelation function (rotation + x + y). If * is answered the crosscorrelation function will not be saved.]

.File numbers:1-10,12-100,200-1000
[Enter numbers of the files that should be aligned. This program tolerates gaps in the file series.]

.Optional single ccf output file or *: Singlecc001
[Enter the name to receive a single ccf between two Radon transforms. This is the radon cross-correlation that is inverted for a slice in the 3D CCF. This option is mainly intended for debugging unusual results.]

If a name was answered:
.Enter the ang. step number for the CCF: 10
[enter at which angular step the single ccf should be saved.]

.Search angles, from, to, increment: -180,170,10
[Enter the agular search range and step size.]

.Maximum shift in pixels: 12
[Enter the maximum shift. This is a limit for X and for Y (not radius).]

.Enter variable to receive output: v.radali
[Enter a variable name for a named tuple that will contain the correlation values of the first file in the series. e.g:
v.radali.cc cross-correlation maximum
v.radali.ang angle to reference
v.radali.xs x-shift
v.radali.ys y-shift
v.radali.cang angle to reference determined from ccf center of gravity
v.radali.cxs x-shift from center of gravity
v.radali.cys y-shift from center of gravity

Remarks: TO APPLY THE ALIGNMENT, USE ROTATION WITH NEGATIVE ANGLE FIRST, FOLLOWED BY SHIFT WITH NEGATIVE VALUES. (this apllies when the reference is entered fist, as the program is intended. For multireference alignment the program can be used the other way around, i.e. first shift positive then rotate positive.)

Programs: radaling2d.f, and many subroutines in
/extras/fstandalone/applications.

Author(s): M.Radermacher, others (for support routines).

radalignto3d 2D Radon transform alignment to 3D Radon transform
<-> 2D Polar Fourier transform to 3D polar Fourier transform

image format: SPIDER

USAGE: radalign2d

.3D Fourier-Radon input reference: refradf001
[Enter the name of the 3D reference Fourier-Radon transform. This reference should be already appropriately filtered. Minimum filter required in "radfour" is filter 8 (=sqrt(r*)). This high-pass filter in applied to both, the single images adn the reference. With this filter the crosscorelation function has the appropriate weights and sharpness. (alternatives of filtering one with r* instead should be tried first. The transform must be in polar coordinates.]

.Optional file name for extracted slices or *: *
[If you want to extract EVERY slice as it is extracted for 3D alignment, enter a sample name here. The number will be counted up by the program. This can be a very large amount of data and is recommended only for test purposes.]

.Optional file name for single CCF output or *: *
[If you want to save indivicual CCfs, then enter a name. Also for test purposes or for tuning very difficult situations.]

If a ccf file name was given:
.Psi angle index for saved CCF: 10
[To reduce the number of files, enter here the psi angle for which the CCFs should be saved (otherwise there would be too many.]

.Optional peak file name or *: peak0001
[Enter a file name where you can save the angular cross-correlation for each projection. This can be helpful.]

.Template for 2D projection input Fourier Radon files: rad2df00001
[Enter a template name for the input 2D Fourier-Radon files. The number at the end of the name is replaced in the program.]

.Selection document file for 2D input series or *: select001
[Enter document file with continuous keys and indicated by 1 the projections that should be aligned and by 0 those that are skipped.]

IF no selection file was given:
.Enter file number range e.g. 1-100: 1-376
[Enter the file numbers by hand if no selection file was provided]

.Exclude missing lines (E): e
[Enter "e" if you want the alignment to exclude the missing lines in the reference from the alignment process. Missing lines are recognized either by the counter at the end of each line in the Radon transform, or, if not present by a threshold of the line sigma.]

If E was answered:
.Min % of lines present in a CCF: 10
[Enter the minimum of lines that must exist for a cross-correlation to be acceptable. If a lower percentage contributes, the cross-correlation is skipped.]

.Threshold of sigma to call line missing: 0.01
[Enter a threshold for sigma below which the line is considered missing.]

.Create debug output: N
[Enter Y if you want a lot of output that tells you details about the program's progress at various stages.]

.SNR estimate of single line in Radon transform: 1.2
[Enter the estimated signal to noise ration of a single line in the 2D Radon transform. This value is used to correct thenormalized CCF for the influenace ofnoise.]

```
        .Print list of excluded lines: N
        [Enter Y if you wnat the list of lines that are considered missing.]

        .Ang. increment of extracted Radon transform (def=2D input): 0
        [If you want a different increment for the extracted Radon transform
        slices, then enter a vlaue here. Otherwise the slice are extracted with
        the same increments as found in the 2D Radon transforms.]

        .Shift alignment or not (S/N-def): S
        [Enter S if you want to include the translational alignment.]

        If S was answered:
                .Max shift radius: 7
                [Enter the maximum shift length. The larger the slower the
                program. Should be large enough though to be able to align
                everybody.]

                .X-shift, Y-shift or Both=def (Y,Y,B): B
                [Enter if you want to check the shift only in a specific
                direction. Usefull when aligning sticks, where a shift alogn the
                stick may not be justified.]

        .Subsearch around header angles(S/N): S
        [Enter S if you like to search around the angles found in the header, N
        if you want to do a full search.]

        .Angle set numbers in file header for reading, writing (def 1,2):1,2
        [Enter where to read the angle for the subsearch and where to write the
        resulting anlges. The spider header has three positions for Euler
        angles. Originally there was also spce for shitfs, which disappeared in
        later versions. Now only position 1 has this space.]

        .Phi search from, to,increment: -25,25,5
        [Enter the search range of the first Euler angle around Z.]

        .Theta search from to incrment: -25,25,5
        [Enter the search range for theta, the second Euler angle around Y']

        .Psi search from, to, increment: -25,25,5
        [Enter the search range for the third Euler angle (in the projection
        plane.)]

        .Automated low-pass filter Y/N : N
        [If y is answered the CCF calcuations are low-pass filteresd according
        to Crowther's formula (d=pi/N*D).]

        if Y was answered:
                .Ratio of object diameter to volume x-dimesion: 0.8
                [Enter the volume diameter as a ration to the X-dimension of the
                volume.]

        .Name for output document file: alidoc001
        [Enter the name of the document file that will receive the new agles and
        shifts. (phi, theta, psi, xs, ys, radial shift, shift direction)

        .Enter varaible to receive output: v.ali
        [Enter a vairable for receiving the alignmen of the first projection.
         It will contain e.g. v.ali.psi, v.ali.theta, v.ali.psi, v.ali.xshift,
         v.ali.yshift.]

Programs: mrrmornia.f, and many subroutines in
          core/fstandalone/applications.

Author(s): M.Radermacher, others (for reading/writing and support routines).
```

radextract extract 2D Radon transform from 3D Radon transform

image format: SPIDER

Purpose: Extract a central section through a 3D Radon transform at arbitrary
angle. The inverse of the extracted transform is theprojection et
the pecified angle.

USAGE: rad extract

.Input 3D Radon file: rad3d001
[Enter input 3D Radon transform. This may also be the Fourier
transform of the Radon transform.]

.Output 2D Radon file: rad2d001
[Enter tha name of the output 2D Radon transform. If the input is
Fourier, then the output is also Fourier.]

.angles phi, theta, psi: 35,60,40
[Enter the angles for the extraction of the radon central section.]

.ang. increment of output transform: 2.
[Enter the angular increment to be use din the 2D output transform.]

Programs: picrad.f, extractradon.f and support routines.
Some programs extracted from SPIDER.

Author(s): M.Radermacher

radfour **Calculate the 1D Fourier transform of a 2D or 3D Radon transform**
or Fourier filter a Radon transform.

image format: SPIDER

USAGE: radfour

.Input file: rad2d001
[Enter the name of the 2D Radon transform.]

.Fourier Output file or *: radfour001
[Enter an output file name for the Fourier Radon transform. Enter
* if you do not need it, e.g. when this program is just used to
Fourier filter the Radon transform.]

.Filtered Radon transform: rad2dff001
[Enter the file name for the Fourier filtered Radon transform or
* is you do not need it.]

.Pad to larger dimension? (P): P
[Enter P if you want to padd the Radon transform before filtration,
anythng else if not (for example N).]

if P was answered:
.padding dimension: 128
[Enter the new p-dimension of the Radon transform.]

.Fourier Filter? (Y/N): Y
[Enter Y if you want to apply a filter, N if not]

if Y was answered:

.Number of filters (max 3): 3
[Enter how many filter you like to apply. The options
below will be repeated for each filter.]

Filter options:
1,2 cutoff low-pass and high pass
3,4 Gaussian low-pass and high-pass
5,6 Fermi low-pass and high-pass
7 r* filter
8 sqrt(r*) filter
9 apply filter as provided as a lin in a file.

.Filter type: 5
[Enter the type of filter you want to apply, in this example
(5) a Fermi low-pass filter.]

For filters 1-7 the next question is:
Radius: 0.2
[Enter Fourier radius. If the value is below/equal 0.5
it is interpreted as an absolute Fourier radius. If it
is larger than 1 it is interpreted as a radius in Fourier
pixels. It is highly encouraged to use the absolute radius.]

For filters 5 and 6:
.Temperature: 0.02
[Enter the temperature fot the Fermi filter.]

For filter 9:
.Input filter file: filter001
[Enter the file that contains the filter function. This file
has standard spider image format, and an x-dimension equal to
the p-dimension of the Fourier Radon transform. Note that the
p-dimension of the Radon transform is 4 less than the Radon
file x-dimension, since the last 4 positions are reserved for
extra information.]

.Which line from the filter-file should be used: 1
[A filter file can contain multiple filters, one per line.
Enter the line number for the filter you want to use.]

.Preserve average (P)

```
 Divide average by the number of rows (D)
 Apply filter also to average (A): A
[Enter how you would like to treat the average of each line,
 which is the 0 Fourier coefficient. The average can be
 preserved, divided by the number of rows in the image or
 can be applied to the average. This can be important for
 Radon inversion.]

.Amplitude normalozation? (Y/N): N
[This option allows to normalize all amplitudes which may be useful for
 phase cross-correlation.]

if Y was ansered:
        .Epsilon for Amp. Norm.: 0.1
        [Enter a value below which the amplitude is not normalized
         anymore to avoid divide by 0 or noise amplification.]

.Sigma normalization? (Y/N): N
[Enter Y if you want to normalize each line in the Fourier Radon trans-
 form with its sigma. Useful for some cross-correlations. The cross-
 correlations in EMIRA for Radon transforms do this normalization
 anyway, so it is not needed here.]
```

Programs: em_radfour.py, fouradcalln.f, fouradn.f, four2.f and subroutines

Author: M. Radermacher (except for four2 and below.)

radmake3d **Create an empty 3D Radon transform (or Fourier-Radon** transform)

image format: SPIDER

USAGE: radmake3d

.Output (empty) 3D Radon transform: 3drad001
[Enter name for the 3D Radon transform]

.Type of Radon transform, Fourier or real (F/R=def): R
[Enter if you want a real space of Fourier space Radon transform. Default is real space. Answer can also be longer, like "real" but only the first letter is used.]

.Enter p-dimension, increments in theta and phi: 128,2.,3.
[Enter the p-dimension and the angular incements in theta and phi. The created volume will have the p-coordinate along x, the phi-coordinate along y and the theta coordinate aling z.]

REAMRKS: The 3D Fourier-Radon transform is essentially a 3D Fourier transform in Polar coordinates. The x-dimension of the file will be 4 pixels longer than p in real space and 6 pixels longer than p in Fourier space. The last four pixels are used to store a projection counter and line-sigmas.A Fourier transform needs 2 (1 complex) number more than the real space line.

PROGRAMS: em-make3drad.py, in: fstandalone/applications: mrrmake.f

AUTHOR(S): M. Radermacher
Author: M. Radermacher

radsigstat Calculate histogram of sigmas in Fourier Radon transforms

image format: SPIDER

PURPOSE: Calculate the histogram of the sigmas found in the Fourier Radon transform. The program was designed to dtermine the value of sigma usable as the treshold for finding the missing data area.

USAGE: radsigstat

.Input Fourier Radon transform: rad3df001
[Enter the name of the image containing transparencies]

.Output document file: sigstatdoc001
[Enter name of output document file. This will contain 3 columns of data:
column 1: value of sigma (histogram interval lower border),
column 2: histogram value (counter)
column 3: cumulative percentage of histogram counts.
If you want to use the document file in SPIDER v.5.0 use new2olddoc to convert it.]

.Number of bins (def. 100): 120
[Enter the number of bins used in the histogram. The program divides the range between the minimum value and maximum value of sigma that is found in the Fourier Radon transform.]

Programs: em_radsigstat.py, radonsigstat.f

Author(s): M. Radermacher

radsum3d **Average 2D Radon transforms into 3D Radon transforms.**
(Also works for Fourier Radon transforms)

image format: SPIDER

USAGE: radsum3d

.Input (empty) 3D Radon transform: rad3d001
[Enter the name of a 3D Radon transform to which the 2D projection transforms should be added (averaged)]

.File name pattern for 2D projection transforms: rad2d00001

.Should the projections be added or substracted: +
[Enter - if the projection transforms ahould be subtracted and + if they should be averaged into the transform. The subtraction can be interesting fot iterative projection alignment.]

.Use equalized sampling or polar sampling: P
[You can shose between polar sampling, and (E) a sampling that equalizes the sampling points by diluting them with the cos of the latitudinal coordinate.]

.Are the radon transforms real or Fourier: R
[Since there is no indication is the file if the Radon transforms are real valued or Fourier transformed it has to be indicated here. The 3D and 3D transforms both must have the same type (Fourier or real)]

.Set line value to 0 when counter is on 0 (Y/N=default): N
[Enter Y if the values in the line should be set to 0 when the counter at the end of the line is on 0, meaning that there are not projection transforms averaged into this line. This can occur when the - option is used to subtract projection transforms. After multiple iterations numerical errors may leave small values in a line, even though they should be 0.]

.Input selection document file or *: sel001
[Optional, enter the name of a selection document file that tells the program which images to use and whic ones to leav out.
*Typically, this file has the file number as key, and in the first column a 1 if the projection should be used and 0 if not.
*Alternatively the first column may contain any value, and the decision if a projection is used or not depends on a threshold (see below). This can be used in alignment procedures where one can take the cross-correlation coefficient as collumn 1 and give a threshold as a criterium for removing projection transforms that show a cross-correlation coefficient below a certain threshold.
*Second alternative: The document file contains a key followed by the image file number, followed by one or more numbers indicating the set membership(s) of the particle. This format was created for the use of new data partitioning methods under development.]

If a selection file was specified:
.Set number to reconstruct: 0.
[Enter the set that should be reconstructed. The document file created by hex3drm has multiple sets.]
if the set number was not 0, then the program asks:
.Do the sets overlap (Y/N = default): Y
[Y means that one particle can belong to more than one set.]

if the set number was 0 (the normal situation):
.Threshold for selection document file (def 0.5): 0.5
[Enter a threshold value for the value from the selection document file. Below this value projections are rejected, above accepted. 0.5 works well for the simple 0 / 1 index. If a 0 is entered other questions follow]

If the threshold is specified as 0:
.Sigma to calculate the selection cutoff (def -0.2): -0.5
[If no threshold was given, the program determined the sigma and average of the selection column. Specifying

```
                0.5 her means that all images with a slection value of
                less than average-0.5*sigma are rejected.]

If no selection file was specified:

        .Enter file numbers (style: 1-100,120-134,137,150-1000): 1-10594
        [Enter the numbers of the files that should be averaged into the
         3D Redon transform. At this time it is not clear if there is
         a limit to the line length or if continuation lines a possible
         (probably not).]

        .Randomization with replacement (Y/N=default): N
        [Enter if you want to do a reconstruction using randomization with
         replacement. This can be used to obtain an approximation of the
         3D variance. Default is N]

         If Y was answered:
                .Number of draws: 9000
               [Enter the number of projections that should be drawn randomly
                from the available projection set.]

               .Name of statistics output doc file: statdoc001
               [This document file will have the information about which
                porjections were drawn.]

        .Angle set number in projection header (1=def, 2 or 3): 1
        [The spider image file contains 3 locations to store euler angle sets
         and x-y shifts (although the shift locations have been used also for
         other purposes in the later versions of SPIDER. Typically the valid
         Euler angles can be found in position 1. The purpose of multiple
         location more redundancy for alignments. When 2D Radon trasnforms
         are aligned to 3D Radon transforms (or 2D polar Fouriers to 3D
         polar Fouriers which is the same), then the new angle can be saved
         in a position other than 1, and a reconstruction can be first
         calculated to possibley verify the validity of the alignment, before
         overwriting  the default position.]

        .Symmetry to impose 0=none=def, 1 around z-axis, 4 icosahedral: 0
        [Enter 0 for asymmetrical particles, 1 for particles with proven
        symmetry around the z-axis, which must be well aligned, and 4 for
        icosahedral symmetry enforcement. Other symmetries are in the
        works.]

Programs: radsum3d.py, put2drads, cputrmn_plus.f, putrmrealn.f, getsym.f

Author: M. Radermacher
```

readdoc read values from a document file

file format: SPIDER document file

USAGE: readdoc

```
        .Input docfile name: doc001
        [Enter the name of the document file you want to read]

        .key followed by vaiables: v.key,v.var1,v.var2  (or *, see below)
        [Enter a key followed by variables that will recieve the values
         that are stored in the file under that key. The values are assigned
         to the variables in the order they occur in the document file. If
         an * is given the program only returns the first and last key in a
         named tuple called "doc", v.doc.minkey and v.doc.maxkey. The
         information is always provided upon the first reading of the
         document file.]
```

Notes: USE readdocclose TO FREE UP THE MEMORY USED BY READDOC. Readdoc reads the document file into memory and only keeps the latest occurence of a key if a key is used multiple times. Within a session the document file is only read upon the first call to readdoc. If the file changes it must be closed before the new content can be accessed with readdoc again.

Programs: adm_master.py adm_readdoc.py

Author(s): M. Radermacher, 2012

readdocclose Close document file to free up memory

file format: SPIDER document file

USAGE:
        readdoc close

        .Enter file name: dofile001

Program: adm_readdoc.py integrated into the core of EMIRA.

remove outliers Remove outliers from document file columns

file format: SPIDER document file

PURPOSE: Remove outliers from columns in a document file. The program was to sort through the coorinates created by correspondence analysis or PCA. Since real outlier images have a major influence on the direction of the eigenvectors, it is important to remove them. Be carefull if you are removing more that 10-20% of your images. You may be removing a complete cluster. Best is to cofirm the removal by checking the factor map for the location of these particles. Since the output file contains the information, which coordinate caused the removal it is easy to look at the corresponding map to visually confirm this.

USAGE: remove outliers

.Input coordinate docfile: imccoord001
[Enter the name of the document file that you want to process.]

.Output selection docfile: removelines001
[Enter name of output document file. This will have the keys of the lines to be removes, followed by a 0, followed by the input column number that was the reason for the removal. Keys may be occuring multiple times, if the exclusion is based on multiple columns. This file can be appended to a typical 0/1 selection document file, or, if such an existing selection file is entered as output name, it will be appended.]

.Output doc file format (0=new,default,1=old): 1
[Enter the format of the output doc file. default is 0 (= new format). option 1 added for compatibility with SPIDER version 5.0].

.Columns to include: 2-4,6
[Enter which columns of theinput document file should be checked for outliers.]

.Sigma multiplier for threshold: 3.3
[Enter the factor by which the standard deviation of a column is multiplied to determine outliers. Outliera are those that have values smaller than average-factor*sigma or larger than average_factor*sigma.]

.Number of columns to write to the output docfile: 3
[The minimum number is 1, which will only write the 0 to the output file. 2 will also wrte the column that caused the out lier. Anything larger will be 0s. The reason to put multiple columns is because the selection file used may have extra information in each line, and adding a shorter line could create problems in reading the file later.]

Programs: em_removeoutliers.py, doceliminate.f

Author(s): M. Radermacher

rot rotate an image

image format: SPIDER

USAGE: rot

```
        .Input file: image001
        [Enter input image name.]

        .Output file: rotated001
        [Enter output image name. Directory will be created if needed.]

        .Angle, xshift, yshift: -30,2.3,1.5
        [Enter rotation angle and x and y shift values. Default 0 for
        all three. Shift is applied after rotation, bilinear interpolation.
        If the rotation angle is 0, than also no shift will be applied,
        independent on input. (This is a feature of rot32)]


Programs: rot32_m.f (compiled to rotate2d), rot32.f in
          /extras/fstandalone/applications.
          Program extracted from SPIDER.

Author(s): M.Radermacher, others.
```

rot3d rotate a volume

```
image format: SPIDER

USAGE:       rot 3d

        .Input file: volume001
        [Enter input image name.]

        .Output file: rotvol001
        [Enter output image name. Directory will be created if needed.]

        .phi,theta,psi: -30,42.3,111.5
        [Enter rotation angles. Euler angles as defined in SPIDER. I.e.
         first phi around z then negative theta around y' then psi around
         z''.]

        .shift center by x,y,z (integer): 0,0,0
        [If 0 or nothing is answered the rotation center is the default
         SPIDER rotation center (n/2+1 in all 3 dimensions). Otherwise
         the values provided are subtracted from the default origin.

        .linear or quadratic interpolation (def. Q): qadratic
        [Enter if linear of quadratic interpolation is to be used. Only
         the first letter is recognized, thus answers 'q' or 'l' are
         enough but full answer is possible.]

Programs: rotas3_m.f, rotas3.f and subroutines.
          Program extracted from SPIDER.

Author(s): M.Radermacher, others.
```

scalevalues Scale density values in an image

image format: SPIDER

USAGE: scale values

  .Input file: image001

  .Output file: scaledimage001
  [Enter output file. Directory will be created if needed.]

  .Enter formula, use P1 as input pixel value: ((p1+v.sc1)*5)
  [Enter the equation the scales each pixel in the image.
  Examples of additional accepted formulas:
  THIS IS NOT FULLY TESTED. ALL EXPRESSIONS MUST START WITH
  (AND END WITH). WORKING IS: (LOG((P1+1)*1)) AND THE EXAMPLE
  ABOVE.
  IN THEORY THE FOLLOWING SHOULD WORK BUT SOME HAV PROBLEMS:
  ((P1-2.)/2.+SIN(P1+2))
  ((P1/512.)+0.)
  (EXP(LOG((P1/512.)+0.)+v.x+v.y))
  (4+ABS((P1-2.)/2.))
  (P1**3+P1+EXP((P1-2.)/2.))
  Any number in the formulae above can be replaced by a variable.

Programs: arith_m.f, in fstandalone/applications.
    Program extracted from SPIDER.

Author(s): M.Radermacher, others.

setangle Put Euler angles in the header of an image.

image format: SPIDER

USAGE: set angle

```
        .Input file: image001
        [Enter the name of the input file.]

        .Position of Angles (1,2 or 3): 1
        [There are three positions for angles and shifts in the image header.

         Note: the set program has been made consistent with current (2013)
         SPIDER header locations. Position 3 is different than in spider 5.0.
         Also, if the Radon alignments in spider 5.0 was used, position 2
         writes the shift reaching into the new position 3. The reason is
         that there has been a change in the SPIDER header, after the
         alignment programs were developed, which created this conflict.
         Originally these positions were created specifically for
         Radon transform alignment. This was ignored in the redefinition
         of the header.]

         .Angles phi,theta, psi: 10,25,-90
         [Enter the three Euler angles (phi around Z, theta around Y' and
          psi around Z'').
```

Author(s): M.Radermacher
Programs: em_setangle.py, putheader.f

shift shift an image or volume

image format: SPIDER

USAGE: shift

```
        .Input file: image001
        [Enter input image name.]

        .Output file: shifted001
        [Enter output image name. Directory will be created if needed.]

        .xshift, yshift, zshift: -3.0,2.3,1.5
        [Enter shift values for x,y, and z. defaults all 0.]

        .Mode, Fourier, Real space: R
        [Enter if Fourier (F) or real space (bilinear) (R)  interpolation should
        be used. Default real.]
```

Programs: shiftf_m.f shiftr.f shift3.f and others. located in
/extras/fstandalone/applications.
Programs extracted from SPIDER.

Author(s): M.Radermacher, others.

sqroot square root of an image

```
file format: spider

USAGE:        sqroot

        .Input file: image001
        [ Enter input file name]

        .Summed Output file: image002
        [Enter name of output file containing the square root of the
         input image. Negative values in the input image are set to
         -sqrt(-value).]

Programs: add.f (extracted from spider)

Author(s): M.Radermacher
```

square square an image

file format: spider

USAGE: square

  .Input file: image001
  [Enter input file name]

  .Output file: image002
  [Enter name of output file that will contain the square of the
   input image.]

Programs: add.f (extracted from spider)

Author(s): M.Radermacher

subtract subtract 2 images

file format: spider

USAGE: subtract

  .Input file (minuend): image001
  [Enter input file name of the minuend image]

  .Input file (subtrahend): image002
  [Enter name file name of the subtrahend image]

  .Output file: image003
  [Enter name for the difference image.]
Programs: add.f (extracted from spider)

Author(s): M.Radermacher

table2doc convert whitespace seprated table to SPIDER document file.

file format: input: text, white space separated columns, output SPIDER doc. file

USAGE: table2doc

```
            Input text file: 1_31_14_WITHDBQBLANK.txt
            [enter input text file]

            Output doc file: tabledoc001
            [enter output file name.]

            Number of columns to read: 2
            [Enter the number of columns you like to read.]
```

Explanations: This program reads the file line by line. If a the first character
in a line is not a number, the line is skipped.

Program: /fstandalone/coltodoc.f, em_table2doc.py,

Author: M. Radermacher, 2013

threshold Theshold an image

```
USAGE:        threshold

        .Input file: image001
        [Enter name of input file]

        .Output file: imagethreh001
        [Enter name of output file]

        .Masktype (S)imple, (T)hreshold, (F)ixup: s
        [Specify the type of thresholding.
        S replaces everything above or below the threshold with the
        corresponding threshold value.
        M create a binary image, 1 above (or below) the threshold, 0
        otherwise.
        F will ask for a value to place where the threshold is exceeded,
        can be used for dust removal.

        For option S or M:
                (A)bove, (B)elow, (C) both sides: c
                [Enter if values above, below or on two sides of two
                threshold values should be replaced]

        For option F:
                .Fixup value: 0.1
                [Enter value that should be places in the pixels that
                exceed the value]

        .threshold, or lower upper thresholds:-1.0,1.0
        [Enter either just one threshold value, or two, dependent on which
        options you used.]

Programs: threshold.f, em_threshold.py

Author: M.Radermache, threshold.f adapted from thresh.f in SPIDER
```

variables, their use in EMIRA

EMIRAda uses variables. All variables start with v. Since variable processing is directly piped into the Pyhton shell, the following words are not allowed as part of a variable (Python 2.7 keywords):

and	del	from	not	while
as	elif	global	or	with
assert	else	if	pass	yield
break	except	import	print	
class	exec	in	raise	
continue	finally	is	return	
def	for	lambda	try	

For example, the variable v.class creates unpredictable errors.

msavisma Create a visual map of msa results

file format: special and spider

USAGE: msa vismap

```
        .Output file: vismap012
        [Enter the name of the image that will receive the visual
         representation of the map. This is a large image, divided
         into a grid, which each will receive the average image of
         all images whose coordinates fall into this grid square.]

         .Number of divisions: 10,10
         [Enter the number of divisions in x and y to be used to define the
          grid.]

         .Upper, lower contrast limit in sigma units: 4.,4.
         [Define the threshold that is applied to the average images. This
          helps to obtain a well visible map. It takes out spikes maybe
          caused by dust in the images. This threshold should not be
          defined to low, since otherwise the images have a binary appearance
          and important details may not be visible anymore. 4,4 is the
          default and a good value to start with.]

         .Map output document file: visdoc012
         [Enter the name for a document file that will contain a list that
          identifies which images contribute to the average in which grid
          square. The sqares are numberred continuously starting in the upper
          left corner (coodinate 1,1) then going along the x-direction continuing
          in the next line etc.. For example a visual map divided into 10x10
          grid squares will use numbers from 1 to 100.]

         .Imput img. coord. doc file: imcdoc001
         [Enter the document file that contains the image coordinates. This file
          needs to be created with "msa imc2doc" from the output image coordinate
          file of correspondence analysis (prefix_IMC.ext).]

         .First key, last key: 12,19567
         [Enter the first and last key for the document file. Typically first
          and last image number.]

         .2 factor numbers: 1,2
         [Enter the numbers of the two factors that will span the 2D map.]

         .Map standard deviation limit: 5.
         [Enter a limit for the standard deviation of the map. This removes
          outliers from the display. Depending on the distribution of the
          image coordinates, however, it may also exclude areas that may be
          important. For excample a small cluster seprated from the main
          population. Use a large value for a first display, between 5 and 10.]

         .First input file: image00001
         [Enter the name of the first image that is used for avaraging. The
          image must exist. Note: image 1 may not always be part of the
          existing image set.]
```

Programs: vismap.f, em_vismap.py
Author: M. Radermacher

vs2doc **convert image coordinates from an XMIPP neural net to** a spider document file

file format: .vs file (XMIPP, input), SPIDER document file (output)

USAGE:

```
vs2doc

Input XMIPP vs file: alu.vs
[enter input file created by XMIP vsom]

Output doc file: Imcoo001
[enter output file name.]
```

Explanations:

The vsom input file typically has a name of the form alu.vs (three letters followed by ".vs") it contains example:

```
    3 hexa 11 11 gaussian
    10 4 60.9965 ../../alifwi/alifwi00002.yhw
    3 0 69.4382 ../../alifwi/alifwi00003.yhw
    10 7 69.4475 ../../alifwi/alifwi00004.yhw
    8 5 75.8319 ../../alifwi/alifwi00005.yhw
    9 3 71.3168 ../../alifwi/alifwi00006.yhw
    9 2 70.592 ../../alifwi/alifwi00007.yhw
    5 0 69.4099 ../../alifwi/alifwi00008.yhw
    ....etc
```

The first line describes the topology of the neural net. The following lines contain two intergers that describe the node coordinate that a specific image belongs to, followed by a weight, followed by the image name.

The spider output document file looks like this:

```
 ; dlu.vs, 3 hexa 11 11 gaussian
    2 5    10.00000     4.00000   115.00000    55.00000    60.99650
    3 5     3.00000     0.00000    34.00000     4.00000    69.43820
    4 5    10.00000     7.00000   118.00000    88.00000    69.44750
    5 5     8.00000     5.00000    94.00000    64.00000    75.83190
    6 5     9.00000     3.00000   103.00000    43.00000    71.31680
    7 5     9.00000     2.00000   102.00000    32.00000    70.59200
    8 5     5.00000     0.00000    56.00000     6.00000    69.40990
    .... etc.
```

The comment contains the map topology
The key corresponds to the file number
The next two integers specify the square in the 11x11 (see topology) map that contains the image, the next two columns contain the same for a transposed arrangement. When comparing the coordinates between the vsom file and the document file one needs to take into account that vsom starts counting at 0, while spider always starts counting at 1. The last column contains the weight.

Program: em_vs2doc.py, vs2doc.cpp in /cstandalone, compilation in Mvs2doc.

Window **Box out small image or volume from large image or volume**

image format: SPIDER

USAGE: Window

.Large input file: raw001
[Enter name of the larger file from which a smaller images is boxed out]

.Output file: small001
[Enter name of the output image/volume]

.New X-, Y-, and Z-dimensions: 32,32,32
[Enter the dimensions of the output file. If it is a 2D image, only
x and y are required.]

.Upper left corner X-, Y- and Z-coord: 10,25,17
[Enter the large file coordinates of the upper left corner of the boxed
image/volume. The specified point will be the first pixel in the output
file.]

.Backgound value: .12345
[If the output image is not fully contained in the input image, the
missing area will be filled with the background value. Default value
is 0.]

Programs: window_m.f, binary: window; Adapted from SPIDER.

Autor(s): M. Radermacher,

writedoc write numbers into a document file

USAGE: writedoc

```
        .Input docfile name: doc001
        [If the file does not exist, it is created, if it exists new lines
         will be appended. Only the last occurence of a key will be accessible
         for reading (see operation readdoc)]

        .key followed by vaiables: v.key,v.var1,v.var2
        [Enter a key followed by variables. There is no limit to the number
         of variables per line, except it may become difficult to read or
         edit. A negative key can be used to create an empty file. It will
         not be recorded.]

         IF THE DOCUMENT FILE IS USED FOR READ WITHIN THE SAME SESSION IT
         MUST BE CLOSED (see operation writedocclose).
```

Programs: adm_master.py adm_writedoc.py

Author(s): M. Radermacher, 2012

writedocclose Close document file to make is useful for reading.

```
USAGE:
        writedoc close

        .Enter file name: dofile001

Program: adm_writedoc.py integrated into the core of EMIRA.
```

zeiss2spider convert Zeiss/SCAI scanner tif to spider format

image format: SCAI scanner tiff (input), SPIDER (output)

USAGE: shift

```
        .Input file: scan4789.tif
        [Enter input tif image name. The extension tif must be provided.
        for variable file names one can use scan[4].tif (four digit space
        holder).
        IMPORTANT:
        1. Because of the construction of the C-program, the file name must
           have at least 3 characters.
        2. the file name may only be one directory up. (../scans/micro4321.tiff
           works, ../../scans/micro4321.tiff will not work properly and all keys
           in the output document file will be 0)]

        .Output file: converted001
        [Enter output image name. Directory will be created if needed.]

        .output document file: convertdoc001
        [Enter name of output document file. This file will contain the
        micrograph number as key and the values of the histogram cutoffs]
```

Programs: zeiss2spider.c

Author(s): Bill Baxter, M.Radermacher

REMARK: The problem with keys in the document files occurs, because
zeiss2spider.c tries to extract the file number, and checks for the occurence
of '.' to show the end of the number. The crude fix was to start searching
at position 3, which allows now for 1 directory up in the input file name.