EMICOPFTCC.DOC

(last update Feb 19, 1993)

CONTENTS

A. INTRODUCTION

EMICOPFTCC is used after EMICOPFT to determine theta,phi,omega orientation and x,y origin parameters for images of icosahedral particles. This uses cross-correlation procedures to compare "raw" (unfiltered or filtered) image data against a data base of reference projection images produced from a 3D reconstruction or model (e.g. from x-ray crystallography).

EMICOPFTCC is a much faster and more powerful orientation search procedure than EMICOFV, which is a brute-force common-lines procedure. Note however, that EMICOPFTCC CANNOT substitute for EMICOFV if there is no model 3D data from which reference data base can be produced. Thus, EMICOFV is still needed when starting a structure determination for a new particle of unknown structure. A key to the success of the EMICOPFTCC procedure is that comparisons made between the raw and model data can include 100% of the available data, or less if you choose. In common lines procedures, the fraction of available data rapidly drops off as a function of resolution. Thus, in principle, the EMICOPFTCC technique is not resolution limited.

EMICOPFTCC converts each raw image from Cartesian to polar format. Each image is subdivided into a series of annuli (usually each one pixel wide). All annuli are sampled NROT times, with the interpolation performed at a fine enough interval to prevent loss of pixel resolution. In this way, the annuli corresponding to the lowest radii in the particle image are over-sampled by a large amount. A Fourier Transform (PFT) of the image is calculated by Fourier transforming each separate annulus. The PFT therefore provides the Fourier transform of the density distribution at each radius of the particle. This provides a particularly useful representation of the data which allows rapid and sensitive determination of the particle orientation. The program

achieves high speed because it determines the particle parameters in several stages rather than in a single, brute force step. Theta and phi are determined first, with an ambiguity in the sign of phi (corresponding to an ambiguity in the hand of the particle). Next, the sign of phi (relative hand) and the omega angle are determined by a real-space, rotational cross-correlation between the projected model and the image. This correlation is typically performed only for a restricted range of particle radii (IRAD1,IRAD2 are user selected). IRAD2, for example is best set to a value close to or just inside the outer particle edge. Setting IRAD2 too large will add unnecessary background noise and hence reduce the sensitivity of the search procedure. Also, experience shows that values of IRAD1 >> 0 help because the low radii in projected images typically contain a significant contribution from the nucleic acid core (for mature virions) which does not generally carry much signal corresponding to the icosahedral orientation.

Correlations can be computed with Fourier-filtered image data (RES_MIN,RES_MAX) and also for data in which the low-order Fourier Bessel terms are omitted: the Jo terms are usually left out of the calculations because they are so dominant and don't convey orientation information.

Currently, EMICOPFTCC outputs a variety of data including the standard EMICO.DAT data format (in the file EMICOPFTCC.DAT) and statistics on the correlation coefficients as a function of radius (EMICOPFTCC.RADS) and resolution (EMICOPFTCC.RES). The variable, ILIST, is used to designate what output is produced.

B. PROGRAM INPUT

1. INPUT FILENAME FOR PACKED RAW DATA IMAGES (A FORMAT)
2. INPUT FILENAME FOR REFERENCE PFTS (A FORMAT)
3. INPUT FILENAME FOR REFERENCE PROJECTIONS (A FORMAT)
4. J1,R1,R2,RES_MIN,RES_MAX,ILIST,SIG (3I,2F,I,F)
5. IMG_NUM1,IMG_NUM2,NUM_STEP (3I FORMAT)

1. INPUT FILENAME FOR PACKED RAW DATA IMAGES (A FORMAT)

   This specifies the file that contains the raw image data.

2. INPUT FILENAME FOR REFERENCE PFTS (A FORMAT)

   This specifies the file containing the series of reference PFTs produced by EMICOPFT from a 3D model or 3D reconstruction.

.
3. INPUT FILENAME FOR REFERENCE PROJECTIONS (A FORMAT)

   This specifies the file that contains the series of reference projections produced by EMICOPFT from a 3D model or reconstruction.

4. JCUT,IRAD1,IRAD2,RES_MIN,RES_MAX,ILIST,SIG (3I,2F,I,F)

   JCUT specifies the minimum rotational Bessel order (Jn) to include in the calculations. The DEFAULT (= 1) will omit the Jo term (this is recommended for the reasons given above). To include Jo, set JCUT to a NEGATIVE value.

   IRDA1,IRAD2 specify the minimum and maximum radii (annuli) to be included in the calculations. The best way to determine optimum values for these parameters is to run EMICOPFTCC with the initial DEFAULT values (0,NANNULI-1) and ILIST = 1, and check the output file EMICOPFTCC.RADS to see what the correlation coefficients are like as a function of radius. You will typically see a large drop in the AVERAGE correlation coefficient near the outer edge of the particle (IRAD2) and you may also see generally lower correlations at low radii (IRAD1). Thus, IRAD1 and IRAD2 may need some fine tuning to optimize the orientations found by the program.

   RES_MIN,RES_MAX specify the lower and upper resolution limits (in TPU), respectively, of the data to be included in the calculations. These values are roughly the same as those used in EMICOGRAD. Specifying a value of RES_MAX greater than the resolution of the reference data makes no sense and may generate meaningless correlation coefficient values. The DEFAULTS for RES_MIN and RES_MAX are 0.0 and SQRT(2*IDIM**2), where IDIM is the dimension of the Fourier transform. This value for RES_MAX is an ABSOLUTE UPPER LIMIT, and thus is an unrealistic value for real (i.e. noisy) data. Thus, your judgement is MOST IMPORTANT in the selection of this parameter.

   ILIST is a switch used to generate various additional forms of output.

<CR> = EMICOPFTCC.DAT output only.
  1 = EMICOPFTCC.RADS and EMICOPFTCC.RES also output.
  2 = same as above plus asymmetric unit plots of correlation coefficients for each image output to terminal (*.LOG file).

SIG enables the use of a masking function that filters the PFT data on the basis of the variance of the PFT data. PLEASE BEWARE: This option currently DOES NOT WORK CORRECTLY!!! Don't use it!!! SIG specifies the threshold for the variance mask. Set SIG = 0.0 (DEFAULT), to disable the masking option.

5. IMG_NUM1,IMG_NUM2,NUM_STEP (3I FORMAT)

IMG_NUM1 and IMG_NUM2 specify the range of images to test. NUM_STEP (DEFAULT=1) can be used to skip images (e.g. NUM_STEP = 2 will skip over every other image). You may keep entering new values for IMG_NUM1,IMG_NUM2, and NUM_STEP. When IMG_NUM1 is set = 0, or when no input is provided for this variable, the program ends.

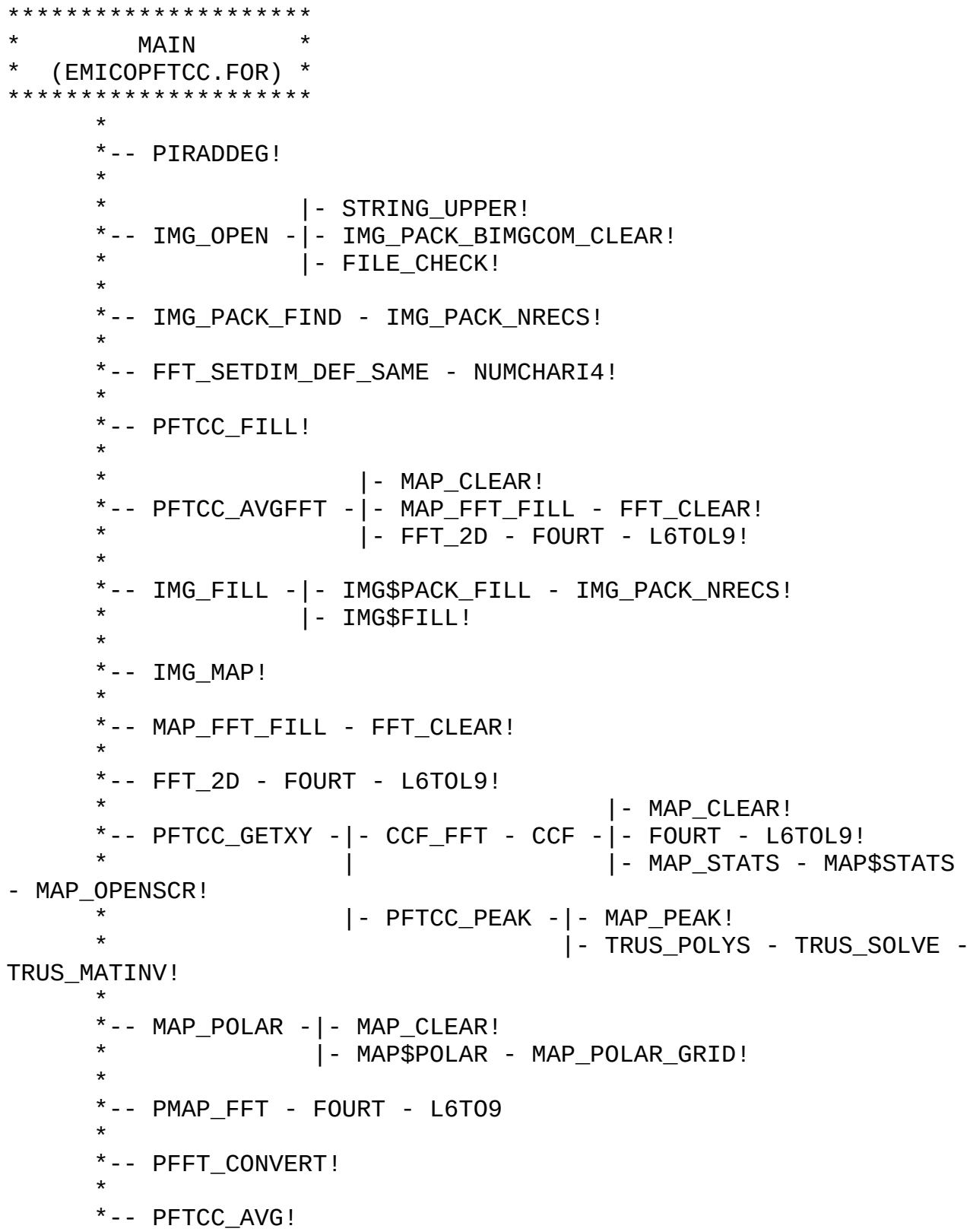
C. EXAMPLE EMICOPFTCC BATCH JOB

```
$ SET DEF JUSTEM$DKA0:[TSB]
$ RUN JUSTEM$DKA0:[TSB.EXE]EMICOPFTCC
BERNAL2:[TSB.VIRUSES.SV40]SV40_TILT0.PCK
BERNAL2:[TSB.VIRUSES.SV40]SV40_TILT0.PFTS
BERNAL2:[TSB.VIRUSES.SV40]SV40_TILT0.PRJS
1, 21, 34, 6.4, 25.5, 1, 0.0
1, 96, 1
$ EXIT
```

D. FINAL NOTES FOR PURDUE USERS

1. SUGGESTION: This program works best when run on JUSTEM or by using the Vaccelerator version ([TSB.AEXE]A_EMICOPFTCC.EXE) on BRAGG.

E. FLOW CHART FOR EMICOPFTCC PROGRAM



```
*
*-- PFTCC_GETTP!
*
*-- PFTCC_GETPRJ!
*
*-- PFTCC_GETPO!
*
*-- MAP_YFLIP!
*
*             |- MAP$ROT - LIB$MOVC3!
*             |- MAP_ROTX!
*-- MAP_ROT -|- MAP_ROTX_SLOW -|- LIB$MOVC3!
*             |- MAP_ROTY!       |- MAP$ROT - LIB$MOVC3!
*             |- MAP_ROTY_SLOW -|- LIB$MOVC3!
*                                |- MAP$ROT - LIB$MOVC3!
*
*-- FFT_2D - FOURT - L6TOL9!
*
*-- CC_RES!
*
*-- FFT_FLT -|- FFT_HIP!
*             |- FFT_LOP!
*
*-- FFT_2DBT -|- COPY_R4!
*              |- FOURT - L6TOL9!
*-- CC_RAD!
```