

EMCORORG.DOC

(last update Sep 3, 1990)

CONTENTS

- A. INTRODUCTION
- B. PROGRAM MODES
- C. EMCORORG PROCEDURE
- D. AUTOMATIC VS. USER-CONTROLLED PROCESSING
- E. INPUT FOR INTERACTIVE JOBS
- F. PROGRAM OPERATION
- G. BATCH-MODE PROGRAM OPERATION
- H. FLOW CHART FOR EMCORORG PROGRAM
- I. FILES REQUIRED TO SETUP EMCORORG PROGRAM

A. INTRODUCTION

EMCORORG uses cross-correlation procedures to determine the location of the origin (center) of particles which appear circular in projection. This program can be implemented in one of three distinct modes:

B. PROGRAM MODES

OPTION 1: IMAGE(S) VS. A SINGLE REFERENCE

A particle IMAGE (or several IMAGES) may be cross-correlated with a single reference particle (which can be either IMAGE or MAP format data). The reference data is typically from a model or circularly averaged data. This option is especially useful for determining initial estimates of centers for icosahedral particles prior to running EMICOFV to determine initial orientation estimates. Results from this procedure are usually superior to those obtained in OPTION #2 (below) because the reference data (e.g. a straight average or a circular average) has higher signal-to-noise than the particle being tested, and because the center of the reference can be determined fairly accurately. This is because the use of a "model" IMAGE (such as that obtained from a rotationally symmetrized correlation average) facilitates the identification of the center of the particle in question, since the center of the "model" is known.

OPTION 2: IMAGE VS. ITSELF ROTATED BY 180 DEGREES

The particle IMAGE is cross-correlated with itself, rotated by 180 degrees. The program identifies the centrosymmetric origin of the particle, which, for particles that appear circular when

viewed in projection, should approximately coincide with the center of the particle. This method is very sensitive to the high noise level in the individual IMAGE. Judicious choice of RES_MIN and RES_MAX (see below) can help, but use of OPTION #1 usually seems to give better results.

OPTION 3: IMAGE(S) VS. CORRESPONDING REFERENCES (BATCH MODE ONLY)

Particle IMAGES are cross-correlated with corresponding references (IMAGE or MAP format data). This option was designed to facilitate accurate refinement of particle origins in the processing of icosahedral particle data. It is run after a 3D reconstruction has been computed (EMICOFB) and after the 3D MAP is reprojected in the corresponding view directions for all the particles in the 3D data set. EMICO is used to produce the BATCH command procedure that computes the set of projections. After this is done, EMICO is also used to produce the BATCH command procedure that executes OPTION 3.

NOTE: OPTION #3 ONLY works as a BATCH job procedure.

EMCORORG is commonly used as a means to determine initial particle origins in the processing of images of spherical virus particles. Further refinement of the particle origins is accomplished with the programs EMICOORG and EMICOORG1 (in conjunction with EMICOFV, SIMPLEX, and GRADIENT) before computing a 3D reconstruction. After a believable 3D reconstruction becomes available, EMCORORG option #3 provides an extremely powerful and accurate way to refine particle origins.

WARNING: Be careful if you set RES_MIN > 0.0 since low resolution information can be crucial for obtaining an accurate estimate of particle centers. Even though the RES_MAX used in EMCORORG can be the same as that used in SIMPLEX(GRADIENT), the use of RES_MIN in EMCORORG is somewhat different than that in SIMPLEX(GRADIENT).

D. AUTOMATIC VS. USER-CONTROLLED PROCESSING

This program provides two methods for identifying the peak position in the cross-correlation function computed. The peak position may be located automatically using a quadratic fit routine kindly provided by Dr. Benes Trus, N.I.H., or with user-controlled interactive graphics. With input data from byte-packed format files (containing "stacks" of IMAGES in a single file), the automatic procedure can be much more efficient than working individually with each particle. With experience you will find

that the AUTOMATIC procedure works quite well and is reliable (i.e. it makes just as good or better decisions than us humans!).

E. INPUT FOR INTERACTIVE JOBS

1. OPTION (I FORMAT)

The user is prompted with the following message:

```
*****
* EMCORORG OPTIONS: (CC = cross-correlate) *
* ----- *
* CC image(s) with ONE reference (e.g. an average) (1) *
* CC image with itself, rotated by 180 degrees (2) *
* CC image(s) with corresponding reference(s) (3) *
* EXIT PROGRAM (4) *
* *
* NOTE: Option #3 ONLY WORKS in BATCH mode. *
*****
```

Note that the only options available in interactive mode are 1, 2, and 4.

2. Enter <CR> or 1 to specify AUTOMATIC or USER-DEFINED peak identification mode (see description above).

3. STORE REFINED ORIGINS IN EMCORORG.TMP? Y/<N>

Touch "Y" to create an ASCII output file named EMCORORG.TMP in which the values obtained for the IMAGE centers (FFT_ORIGX,FFT_ORIGY) are stored. Touching "N" or any other key except "Y" will bypass this option.

4. FILENAME (A FORMAT)

Enter the name of the file that contains the IMAGE(s) whose center(s) you wish to locate.

4a. NUM1,NUM2,STEP (3I)

If the input file contains IMAGE data stored in BYTE-PACKED format, specify the numbers of the first (NUM1) and last (NUM2) IMAGES you wish to examine and the increment (STEP) between images. The default values are NUM1 = 1, NUM2 = number of last IMAGE in the file (BIMG_NUMLIM), STEP = 1. Thus, with BYTE-

PACKED data, you can rapidly scan through an entire (or partial) set of data. With BYTE-PACKED data the program automatically processes successive IMAGES so the user need not retype the filename.

4b. COMPUTE AN AVERAGE? (STORED IN EMCORORG.AVG) Y/<N>
(asked if NUM2>NUM1)

Touch "Y" to create an ASCII output file named EMCORORG.AVG in which a global average of the IMAGES examined is stored. The average is computed with each of the component IMAGES phase shifted to a common origin (for OPTION #1 this will correspond to the center of the reference particle; for OPTION #2 this corresponds to the center of the 2D data array). Touching "N" or any other key except "Y" will bypass this option.

5. RES_MIN,RES_MAX (2F FORMAT)

RES_MIN,RES_MAX set resolution limits for the correlation function calculation. These parameters are expressed in transform pixel units (TPU). The DEFAULT for RES_MIN is 0.0 (i.e. to include ALL low-resolution information). The DEFAULT for RES_MAX = $\text{SQRT} [(\text{IDIM1}/2)^2 + (\text{IDIM2}/2)^2]$. IDIM1 and IDIM2 are the dimensions of the smallest power two transform that can accommodate the IMAGE data. As an example, for a 99 by 99 pixel IMAGE, the smallest FFT would be 128 by 128 and RES_MAX would be set = 90.5. This DEFAULT is the MAXIMUM value that RES_MAX can be set. It is GENERALLY NOT WISE to use the DEFAULT value since the resulting correlation function will be influenced by the inclusion of high resolution noise. Normally, RES_MAX is set < IDIM/2 since the original image SHOULD HAVE BEEN densitometered at a pixel size that is at least TWICE as fine as the maximum expected data resolution. Any time RES_MIN or RES_MAX are changed the new values are used as new DEFAULT values for subsequent calculations. (NOTE: When processing multiple data from a BYTE-PACKED array, the values for RES_MIN and RES_MAX remain unchanged after they are set initially).

6. FILENAME (A FORMAT) (relevant ONLY if OPTION #1 is chosen)

Enter the name of the file containing the REFERENCE data. This file may contain data in either IMAGE or MAP format.

6b. X_SHIFT, Y_SHIFT (2F FORMAT) (DEFAULT = 0.0,0.0)

If the origin of the reference IMAGE/MAP is exactly in the center of the data array, enter a correction factor which

will be applied to the final origin positions that are calculated.

F. PROGRAM OPERATION

A. Terminal display of the peak of the cross-correlation function (CCF)

The program computes the CCF and locates the peak value and its position. The CCF is scaled to normalize the peak value to 255. The values within a 15x15 pixel window, centered on the peak position, are displayed at the user terminal.

The program next calculates by interpolation the exact (floating point) position of the peak (PEAKX,PEAKY) in the CCF. The origin of the CCF is considered to be located at position 0.0,0.0. The shift vector (DX,DY) defines the location of the particle center with respect to the center of original IMAGE data array. The center of the particle is obtained by adding DX and DY to the geometric center of the particle IMAGE array (in pixel coordinates). As an example, for particle IMAGES boxed with circular boundaries using option "C" of EMIMGBOX, the array of pixels is square and odd-sized (e.g. NCOL=NROW=99). Thus, the geometric center of the BOX is at position (NCOL+1)/2,(NROW+1)/2 (or 50.0,50.0 if NCOL=NROW=99).

B. Defining the peak center with interactive graphics

The graphics data tablet puck is used to position the cursor over the center of the peak identified in the CCF. Depress the YELLOW button to list the current values for PEAKX,PEAKY,DX, and DY at the terminal. Values obtained by Trus' quadratic fit method are also given for comparison with the user-obtained positions. Depress the GREEN puck button to EXIT when you are satisfied.

A choice of several cursor types and colors may be selected with a sub-menu by depressing the BLUE puck button. Follow the directions given on the terminal screen. Depress the WHITE button if you wish to further optimize the display of the CCF peak by selecting new values for the number of contour intervals (NCONTUR) used and the density value at which the first contour level (ICLIM) is displayed. The initial DEFAULT value for NCONTUR (the number of contour levels) is 25 (maximum value = NCONTUR_LIM, defined by a PARAMETER STATEMENT in the FORTRAN program [currently = 50]). The initial DEFAULT value for ICLIM is 105. Optimal values for NCONTUR and ICLIM are obtained by trial and error.

You may EXIT the program (1) or continue <CR> with another IMAGE. (For multiple IMAGES from a BYTE-PACKED IMAGE file this is handled automatically).

G. BATCH-MODE PROGRAM OPERATION

All three options of the program can be run in BATCH mode. The simplest way to set up the BATCH command procedures is to use the appropriate sub-menu option provided by selecting option "B" of the main menu in the EMICO program.

A couple notes:

1. When run in BATCH mode, options #1 and #2 require IMAGE input from BYTE-PACKED data format files.
2. When run in BATCH mode, options #1 and #2 automatically send output to the EMCORORG.TMP file and also the EMCORORG.AVG file if more than one IMAGE is processed (the typical case). Thus, if you run EMCORORG as a BATCH process, please be sure to delete unwanted EMCORORG.TMP and EMCORORG.AVG files after you are done.

```
=====
The FORTRAN code for EMCORORG is in [TSB.FOR]EMCORORG.FOR.
This documentation is in [TSB.DOC]EMCORORG.DOC          3-Sep-90
-----
```

H. FLOW CHART FOR EMCORORG PROGRAM

```
*****
*           *
*   MAIN   *
* (EMCORORG) *
*****
*
*           |-- STRING_UPPER
*-- SYS$INTERACTIVE |-- IMG_OPEN --|-- FILE_CHECK
*
*           |-- IMG_PACK_GETRANGE
*-- LX_INIT         |-- GET_YESNO
*
*           |-- IMG_PACK_FIND -- IMG_PACK_NRECS
*-- LX_TABINIT     |-- IMG_PACK_FILL
*
*           |-- IMG_FILL
*-- LX_CURON       |-- FFT_SETDIM_DEF_SAME
*
*           |-- STRING_UPPER
*-- GET_YESNO     |-- FILE_CHECK      |-- STRING_UPPER
*
*           |-- MAP_OPEN -----|-- FILE_CHECK
*-- CORORG12 -----|-- MAP_FILL_2D
*
*           |-- IMG_FFT_FILL -- FFT_CLEAR
```

```

*          | -- FFT_2D -- FOURS -- L6T09
*          | -- MAP_FFT_FILL
*          | -- COPY_R4
*          | -- FFT_FLIP
*          | -- FFT_PSHIFT      | -- FOURS -- L6T09
*          | -- CCF_FFT ----- | -- CCF -- (#1)
*          | -- MAP_SCALE
*          | -- MAP_PEAKSEARCH  | -- LX_ERASE
*          | -- PEAK_SORT_HI    | -- LX_CTSET
*          | -- PEAK_CHK        | -- MAPDSP_VECPLT -
- (#2)
*          | -- CORORG12_PEAK -- | -- TRUS_POLYS --
(#3)
*          | -- IMG_SHIFT        | -- CONTUR_A --
(#4)
*          | -- IMG_ADDIMG      | -- LX_TABPOS
*          |                   | -- CORORG_PLOTINFO
*          |                   | -- CORORG_PLOTINFO
*          |
*          | -- IMG_PACK_FIND -- | -- IMG_PACK_NRECS
*          | -- IMG_DIVIDE
*          | -- IMG_STATS
*          | -- IMG_BKGND
*          | -- IMG_WRITE
*          | -- LX_TABOFF
*          | -- LX_CUROFF
*          | -- LX_CLOSE
*
*          |                   | -- STRING_UPPER
*          | -- IMG_OPEN --    | -- FILE_CHECK
*          | -- IMG_PACK_FIND -- | -- IMG_PACK_NRECS
*          | -- IMG_PACK_FILL
*          | -- IMG_FILL
*          | -- FFT_SETDIM_DEF_SAME
*          | -- IMG_FFT_FILL -- | -- FFT_CLEAR
*          | -- FFT_2D -- FOURS -- | -- L6T09
*          | -- STRING_UPPER
*          | -- FILE_CHECK      | -- STRING_UPPER
* -- CORORG3 ----- | -- MAP_OPEN ----- | -- FILE_CHECK
*          | -- MAP_FILL_2D
*          | -- MAP_FFT_FILL
*          | -- CCF_FFT -- CCF -- | -- (#1)
*          | -- MAP_SCALE
*          | -- MAP_PEAKSEARCH
*          | -- PEAK_SORT_HI
*          | -- PEAK_CHK
*          | -- CORORG3_PEAK -- | -- TRUS_POLYS -- (#3)

```

```
*          | -- MAP_CLEAR
#1  *-- CCF --| -- FOURT -- L6TO9
*          | -- BLANKL
*          | -- MAP_STATS

*          | -- LX_VEC
#2  *-- MAPDSP_VECPLT --| -- LX_VECCLIP

#3  *-- TRUS_POLYS -- TRUS_SOLVE -- TRUS_MATINV

#4  *-- CONTUR_A -- CONTUR_C -- CONTUR_D -- (#1)
```

I. FILES REQUIRED TO SETUP EMCORORG PROGRAM

```
LXLIB.DOC
EMPROGS.DOC
EMCORORG.FOR
EMCORORG.DOC
CONTUR.SUBS
CORLIB.SUBS
FFTLIB1.SUBS
FFTLIB2.SUBS
IMGLIB.SUBS
IMG_PACK.SUBS
MAPLIB.SUBS
MISCLIB.SUBS
PEAKLIB.SUBS
TRUSLIB.SUBS
```