EMCORAVG.DOC

(last update Jul 21, 1987)

CONTENTS

A. INTRODUCTION

ORAVG computes an averaged IMAGE using real-space procedures. The cross-correlation pattern (DEFAULT name CCOR.IMG) computed using EMCOR is examined by EMCORAVG to locate regions in a MAIN (unfiltered) IMAGE that best correlate with a REFERENCE IMAGE. The program provides flexibility in selecting optimized parameters to identify areas in the unfiltered IMAGE which are to be used to compute the real-space average. The MAIN IMAGE is generally a raw, unfiltered IMAGE and the REFERENCE IMAGE is usually either a small portion of the MAIN IMAGE or a portion from a MAIN IMAGE which has been pre-filtered, for example by Fourier processing techniques.

The REAL-SPACE AVERAGE may be computed using two methods, which, when used together, produce optimal results. A brief description of each of the two methods is given below, followed by specific instructions on the interactive use of the program.

B. COMPUTING THE REAL SPACE AVERAGE

METHOD 1: AVERAGING WITHOUT UNBENDING

In this method the positions of the highest correlation peaks are identified and rectangular windows from the corresponding locations in the MAIN IMAGE are summed to compute an average image. The number of windows summed is set by the user by thresholding the correlation pattern to identify the strongest areas of correlation between MAIN and REFERENCE IMAGES. Windows from the MAIN IMAGE are averaged in order of decreasing strength of the correlation peaks. This method of averaging smears out local distortions of the individual unit cells in the crystalline lattice. Thus, this method is best used for reasonably coherent lattice structures or as a preliminary cycle producing an initial

filtered image to be used as reference for further cycles of real-space averaging.

In typical applications of this program, for example with a 512 x 512 pixel MAIN IMAGE in which the unit cells are about 20 x 20 pixels in size, a 64 x 64 REFERENCE IMAGE is boxed from a central region of the MAIN IMAGE. After computing the cross-correlation pattern between the MAIN and REFERENCE images, EMCORAVG is used to compute an initial real-space average (64 x 64 pixels), summing windows from the MAIN IMAGE at the 50-100 highest correlation peaks. This averaged image, when used as a new REFERENCE, produces a correlation pattern with stronger and more sharply defined peaks than were found in the original pattern. The location of peaks will also be less biased by the initial choice of REFERENCE IMAGE. The second average image can be used in another round of averaging, but further rounds are generally unnecessary. Another helpful procedure is to reduce the size of the window after each cycle of averaging to allow more of the local distortions in the crystal lattice to be followed. With noisy image data, if the initial REFERENCE IMAGE is too small (covering an area of only a few unit cells), the correlation pattern may be too noisy to successfully interpret.

With very noisy data it is sometimes helpful to pre-filter the MAIN IMAGE using standard Fourier filtering techniques (EMFFTFLT) in order to obtain a better, initial REFERENCE IMAGE. This, when cross-correlated with the unfiltered MAIN IMAGE, will give a cleaner and more easily interpreted correlation pattern.

METHOD 2: AVERAGING WITH UNBENDING

This procedure removes or compensates for localized distortions in individual unit cells in the crystal lattice. EMCORAVG calculates the average unit cell dimensions based upon the user's initial determination of the principal A and ???? B unit cell vectors using the raster graphics display. Using this initial determination, EMCORAVG then refines the real-space lattice parameters using the cross-correlation peaks themselves. After this, individual unit cells are excised from the MAIN IMAGE file and are re-interpolated into the average or ideal unit cell dimensions. For this procedure to work properly, there must be a large enough number of strong cross-correlation peaks present such that all four unit cell corners are identified in a suitable number of individual unit cells.

C. RUNNING THE EMCORAVG PROGRAM

The program begins by reading in the cross-correlation pattern (DEFAULT name given by EMCOR is CCOR.IMG). Positions of the highest peaks in the correlation pattern are used to identify the best areas in the MAIN IMAGE for averaging. A threshold value (LOWLIM) is used to screen for peaks in the correlation pattern (pixels with values greater than LOWLIM). A value of LOWLIM may be entered interactively and changed until a suitable number of pixels are found > LOWLIM. At this point it must be stressed that the number of pixels surviving this test is generally much greater than the number of "independent peaks" in the correlation pattern since each "peak" usually includes several pixels with intensities > LOWLIM. The pixel data are reduced to a set of discrete peaks by specifying a window (radius) within which there can be only one correlation peak. In general, the radius of the window should be no larger than one-half of the smallest lattice dimension (in pixels) of the crystalline specimen. EMCORAVG sorts all pixels > LOWLIM and then tests each pixel (in decreasing order of pixel intensity) to eliminate all but a single pixel within any single window. The remaining "peaks" identify integer positions in the correlation pattern. EMCORAVG then determines a more accurate (floating point) estimate of the position for each peak by locating the center-of-gravity of a 3x3 pixel window centered on each "peak" pixel in the pattern.

It may be helpful to determine the value of LOWLIM by choosing the option, which interactively uses the raster graphics display of the correlation pattern. The strategy is to use the keypad TOUCH-KEY controls to raise or lower the value of LOWLIM while pixels in the correlation pattern < LOWLIM are blanked out. With some experience it is easy to judge the level at which LOWLIM allows a suitable number of correlation peaks to be identified. If LOWLIM is set too low, the number of pixel values > LOWLIM will exceed program limits (currently 1500 pixels can be handled). Several other TOUCH-KEY options, whose function ought to be obvious, are available to help set LOWLIM.

After LOWLIM is set, the number of distinct peaks must be determined. Specify a distance (RADIUS, in pixels) to define what the minimum separation ought to be between peaks. The program checks, in descending order, the list of pixels with values > LOWLIM and flags those pixels which are closer than RADIUS to another pixel of equal or greater intensity. The number of "peak" positions (NPEAK) in the correlation pattern is then listed. The value for RADIUS should at least be a few pixels to assure that multiple pixels are not found for a single "peak". However, RADIUS shouldn't be made too large (i.e. greater than the smaller of the two unit cell lattice spacings) otherwise adjacent unit cells in the lattice can't be averaged. The best way to judge

whether LOWLIM and RADIUS are correct is to use the option that displays graphically the chosen peak positions in the correlation pattern on the TV screen.

At this point the user has the choice to compute the average image using either METHOD 1 or METHOD 2 as outlined below.

D. COMPUTING THE AVERAGE IMAGE

METHOD1: AVERAGING WITHOUT UNBENDING

After LOWLIM and RADIUS are set, enter the name of the REFERENCE IMAGE file to allow the program to set the size of the averaged IMAGE (can't be > 128 x 128). Note that the REFERENCE IMAGE is not actually used in the program except to determine a DEFAULT size for the averaging window. The user has the option to increase or decrease the size of the window (and final average).

Enter the filename of the MAIN IMAGE from which the average will be obtained (currently limited to 512 x 512 or smaller). The filtered IMAGE is computed by summing areas from the MAIN IMAGE from locations identified by the peak positions. The areas are summed in decreasing order according to peak strength. Thus, the first few areas averaged are those that correlates best with the REFERENCE IMAGE. You may average as many areas as you wish up to the total number of peaks found. You may also produce independent sums of the "odd" and "even" numbered areas to make statistical comparisons of independent averages and determine the level of reproducible detail (resolution) in the data.

If you wish, you may choose the option to interactively display the intermediate results on the graphics screen as the program sequentially builds up the average image. Both the current average and the current area being added are displayed to the left and right of the center of the graphics screen. This option considerably slows execution time, although it can be informative and useful for debugging purposes.

METHOD 2: AVERAGING WITH UNBENDING

After LOWLIM and RADIUS are set, enter the approximate lattice coordinates either interactively with the graphics tablet cursor or by identifying the appropriate file containing them. When the graphics option is chosen, use the tablet cursor to select two diagonally opposite corners outside the boundary of the correlation pattern displayed in the center of the graphics screen (e.g. at 640,512 on the 1280x1024 LEXIDATA graphics device). Note: in order for this option to work properly, the correlation

pattern MUST be displayed in the center of the screen. Now choose a correlation peak with the tablet puck as an origin point for the real space lattice. At the terminal, enter the indices for two peak positions to determine the A and B vectors. Finally, you must mark the locations of the two spots designated in order to generate the lattice. Upon doing this, you may either exit using the GREEN button, or you may continue to mark the A and B spots as you wish. When the lattice is satisfactorily defined, you will be prompted for an OUTPUT file to store these coordinates so that they may be entered using the data file option instead of redefining them constantly. If the data file option is used, simply enter the name of the file of a previously created lattice. Next, you will be prompted for the filename of the MAIN IMAGE data. This will be the image data that will be used in the summation process.

Refinement of the lattice follows. This procedure is done to calculate the average unit cell dimensions. You should enter a percent error margin which will be used to determine what error will be allowed in the calculation of the average unit cell. The first percentage error displayed designates how distorted the actual unit cells were with respect to the coordinates entered by the user. The second percentage error display designates the final distortion of the actual unit cells from the calculated average unit cell dimensions. Then, you should enter the magnification desired for the unit cells. Either the natural magnification may be kept <CR> or any other magnification may be specified.

The next section of the program asks you to specify how the data should be analyzed. The calculated and refined image may be "forced" to the following specifications:

(1) The image produced is such that every unit cell falls upon exact pixel locations so that there are no "interpolation errors" and that every unit cell looks identical to all others. Note that this may slightly distort the image from its real configuration. This is a FORCED REAL SPACE IMAGE.

(2) The image produced is such that its calculated FFT falls upon exact pixel locations. Note that this also may slightly distort the image from its real configuration, and that every unit cell will deviate slightly from all the others due to "interpolation errors". This is FORCED RECIPROCAL SPACE IMAGE.

(3) The image produced is unforced and left in its native

configuration. Note that each unit cell will also be slightly different than all the others due to "interpolation errors". This is an UNFORCED IMAGE.

Upon specifying the type of image to be produced, the resulting modified coordinates of the lattice will be stored in the datafile EMCORAVG.DAT. After displaying the total number of unit cells to be summed, enter the START, END, and INCREMENT of the series of unit cells to be summed. After this is done, decide from the following options how you wish the summation process to be displayed:

NO <CR>:

This option allows the program to sum at its fastest rate and no graphics display concerning the summation process is shown.

FULL IMAGE <1>:

This option displays the entire MAIN IMAGE and identifies the individual unit cells in the order in which they are summed.

UNIT CELL <2> :

This option displays the boundaries of each unit cell magnified on the graphics screen. This option is mainly used to show how much distortion is present as the unit cells are selected and summed. This option will display the average unit cell in bold outline on the graphics screen and then as each unit cell is "picked off and summed", it will display the outline (fainter) of these unit cells. Thus you can easily see how distorted the unit cells are.

E. NOTES

It is important to realize that if this option is chosen, it will only display the naturally occurring distortions if the image produced is UNFORCED!! If any FORCED image summations are displayed, then the average unit cell shown is also FORCED and thus the distortions will not be accurately depicted.

If option #2 is selected, you will first be shown the length and width (in pixels) of the average unit cell. Then you will first be asked to enter a scale factor. This scale factor will magnify the outline of the unit cells so that they can more easily be seen (format F). Then you will be prompted for the radius of

the unit cell points themselves (format I). These two specifications are purely for cosmetic reasons and will enhance the display of the unit cells. After each unit cell is displayed, you must hit the =SPACE= bar to continue with the next image.

After the display options are selected, you will be prompted whether or not the cells are to be summed with different weights. If you wish to do a weighted summation, the program will weight the unit cells according to their distortion from the average cell. Enter the weight to be applied to the cells with the largest distortion.

After the summation is complete, you have the choice of saving the averaged UNIT CELL as an IMAGE file or creating a larger "crystal" by splicing together several cells and saving this as an IMAGE.

==========

The FORTRAN code for EMCORAVG is in [TSB.FOR]EMCORAVG.
This documentation is in [TSB.DOC]EMCORAVG.DOC 21-Jul-87
