Chapter 5

Raw Image Registration (*CROSS-CORR*)

The NEWSIPS system relies upon an explicit determination of the registration between the raw science image and the raw space Intensity Transfer Function (ITF) to apply an accurate photometric correction (discussion of the ITF is deferred until Chapter 6). The image registration is derived using a linear cross-correlation algorithm with the fixed pattern found in all *IUE* raw images as the registration fiducial. In this context the fixed pattern is a manifestation of systematic variations in sensitivity from pixel-to-pixel over the entire image.

5.1 Registration Fiducial

Although the reseaux have traditionally been used for *CROSS-CORR* purposes, using the fixed pattern as the registration fiducial has several advantages over the reseaux; one being that fiducials can be found anywhere on the exposed part of the image, so that many more such fiducials are available than reseaux. Also, the fixed pattern can be detected at even the lowest exposure levels, which eliminates the need for predicted information. Finally, the presence of the exposed spectrum can be excluded from the cross-correlation region, and therefore, does not inhibit the use of the fixed pattern in the surrounding background (De La Peña et al. 1989, 1990). A consistent fixed pattern has been shown to be present in the literally thousands of raw *IUE* images tested; however, it is known that the character of the fixed pattern is modulated to some degree by observation date and camera temperature (Bushouse 1991b).

5.2 General Method

CROSS-CORR employs a pattern-matching technique which depends heavily upon the linear cross-correlation function to identify similar patterns between a raw science image and the

raw-space ITF. A rectangular grid of fiducial positions (centered 28 pixels apart) has been defined for each camera and dispersion. At each defined position, a cross correlation is performed to determine the displacement between the raw image and the ITF at that locale. Specifically, at each locale an array (*template*) of 23×23 pixels of the raw science image is used for comparison to a corresponding array (*window*) of 29×29 pixels of the raw-space ITF image. The window is larger than the template in both dimensions such that the template can be moved about the window to "search" for the best correlation location. It is known from a number of analyses that the maximum extent to the motion of the fixed pattern in most raw images is only a few pixels. Therefore, the window size is ordinarily set in the NEWSIPS processing to accommodate motion of only a few pixels, but the size can be varied, if necessary.

Once a pattern match location is found to whole pixel accuracy at a defined fiducial locale, the match location is evaluated via a series of statistical tests to validate the result. Upon validation, further computations refine this determination to the sub-pixel accuracy required for an accurate photometric correction.

After all the valid pattern match locations for a camera and dispersion have been derived to sub-pixel accuracy for the raw science image, the determinations are viewed as spatial deviations and examined as an ensemble for consistency. Displacements which do not conform in a general sense to a smoothly varying function are adjusted to conform. Finally, after all the fine tuning has been completed, a full three-dimensional (3-D) displacement image ($768 \times 768 \times 2$) is created for ΔSample and ΔLine using a bi-cubic spline interpolation scheme.

5.3   Pattern Matching Algorithm – Step by Step

The following elaborates the procedures incorporated into *CROSS-CORR*. Fiducial locations which have been pre-defined for each camera and dispersion are accessed; these are hereafter referred to as the sampling locations. For each sampling location, the following steps are performed.

1. The raw science image template is classified by its computed median Data Number (DN), and the ITF level possessing a comparable DN value at that location is chosen for correlation purposes.

2. Any reseaux, hot spots, cosmic ray hits, and spectral information are ignored in the template during the correlation analysis. These artifacts contaminate the science template with features not present in the ITF window and are temporarily masked during the correlation procedure.

3. The template is sequentially positioned throughout the window, and at each position a linear cross correlation is performed. For each position a correlation coefficient is computed and the corresponding coordinates (pixel accuracy only) are saved. By virtue of

the sizes of the template and search windows, a total of forty-nine correlation coefficients are computed to determine the best pattern match location to pixel accuracy for each sampling location. These forty-nine coefficients are actually in the form of a 7×7 matrix which corresponds to the overlap positions of the template upon the window. The maximum coefficient in this sample is assumed to be the best representation for a true match.

The cross-correlation coefficient is defined as

$$r(m,n) = \frac{\sum_x \sum_y (w(x,y) - \overline{w(x,y)})(t(x-m,y-n) - \overline{t})}{(\sum_x \sum_y (w(x,y) - \overline{w(x,y)})^2 \sum_x \sum_y (t(x-m,y-n) - \overline{t})^2)^{1/2}} \tag{5.1}$$

where $-1.0 \leq r \leq +1.0$. In this equation, $\overline{t}$ is the average value of the template, $\overline{w(x,y)}$ is the average value of the portion of the window which is currently coincident with the template $t(x,y)$, the line coordinate $m = 0, 1, 2, ..., (L-1)$, and the sample coordinate $n = 0, 1, 2, ..., (S-1)$. For the purposes of this implementation, $L = 29$ and $S = 29$ which are the nominal dimensions of the correlation window; the template dimensions are 23×23 pixels. Summations are performed over the coordinates common to both w and t.

4. The maximum correlation coefficient must pass several statistical tests to be considered a valid correlation:

   - There must be a minimum of 139 (13×13) non-rejected pixels in the template for the correlation to be performed at all (see Item 2);

   - The maximum correlation coefficient cannot reside on the edge of the correlation matrix;

   - The probability distribution criterion of less than 1% must be satisfied;

   - There can be no other "local" maxima in the correlation matrix within 0.25 sigma of the absolute maximum coefficient. Sigma is the standard deviation of the correlation coefficient matrix.

   - The maximum correlation coefficient must be at least two sigma greater than the rest of the coefficients in the matrix.

5. If the coefficient is deemed valid, the position of the pattern match is then determined to sub-pixel accuracy. The original ITF window data is resampled into successively finer and finer sub-pixel increments and the cross correlation with the template is repeated until 0.125 sub-pixel accuracy is achieved.

   If the coefficient is not validated by the tests described in Item 4, the sampling position is considered a bad location, and the algorithm passes onto the next sampling position.

5.3.1 Intermediate *CROSS-CORR* Output

The intermediate results from *CROSS-CORR* can be used to monitor the geometric/photometric state of the raw science image. In the vector displacements (VD) file, these intermediate results are stored in a binary table extension containing the following output for each sampling position:

1. science image x-position (sample),

2. science image y-position (line),

3. ITF x-position of best match (sub-pixel accuracy),

4. ITF y-position of best match (sub-pixel accuracy),

5. maximum cross-correlation coefficient (computed for pixel location accuracy),

6. number of pixels used to calculate the correlation, and

7. the ITF level used for *CROSS-CORR*

Items 1 and 2 are the pre-defined fiducial locations about which *CROSS-CORR* performs its search. Items 3 and 4 are the best match positions between the raw science image and the ITF image. Note that these final positions are in quantized units of 1/8 pixel, which is a consequence of the design of *CROSS-CORR*. Item 5 is the maximum correlation coefficient which can range between 0.0 and 1.0. In practice the correlations range between -1.0 and $+1.0$, but negative correlations have been deemed invalid. Item 6 is the number of points used to compute the correlation; this will range from 139–529 pixels for a 23×23 pixel template. Item 7 is the ITF level used for the correction and will range in value from 1 to 12.

The total number of entries in the binary table extension is both camera and dispersion dependent. Table 5.1 contains the number of fiducial locations used in *CROSS-CORR*.

Table 5.1: Number of Fiducial Locations

Camera	Number of Locations Low Dispersion	Number of Locations High Dispersion
LWP	150	490
LWR	151	511
SWP	139	510

The data contained in the binary table extension can be used to produce diagnostic plots for *CROSS-CORR*. Panel *a* of Figure 5.1 shows the displacements computed between an

SWP low-dispersion image and the SWP ITF. The small boxes denote the fiducial locations where a successful correlation was determined. The associated vector represents the computed displacement between the raw science image and the ITF; the vector displacements have been magnified for visual clarity. Single points mark the locations of fiducials with unsuccessful correlations. Panel *b* of Figure 5.1 shows the magnitude of the correlation coefficients corresponding to the results in Panel *a*. Larger plus signs denote stronger correlations. The dashed lines represent the region of the image that is photometrically corrected. Recall that all of these correlations have passed the 1% probability distribution test which means that there is less than a 1% probability any of these correlations could have arisen by random chance. The corresponding example for high-dispersion data is illustrated in Figure 5.2.

5.4   Evaluation of the Raw Correlations

After all of the pre-defined fiducial locations have been investigated as individual entities, the computed results are examined as an ensemble of information. The goal is to create an idealized or smoothed version of the actual displacements and then compare the smoothed values to the actual values, making adjustments to the actual displacements if necessary.

The creation of the smoothed displacement image is done in two major stages. First, an averaging procedure is invoked for each successful correlation sample. The closest fiducial locations to the sample in question which also have successful correlations are weighted according to their "distance" from the sample in question and if there is deemed to be enough data, the displacements from all of these successful locations are averaged to yield a mean displacement for the smoothed image at the sample location in question. If enough data are not acquired in the pre-defined search region, the data point must be an isolated or semi-isolated correlation. Isolated or semi-isolated correlations which cannot adequately be compared to the immediate surroundings for spatial trends are considered invalid correlations after-the-fact and are permanently removed from consideration.

Once the averaged image is created, the displacements for the "missing" fiducial locations are filled in by an interpolation/extrapolation scheme. The search for missing data is performed in a radial sense, starting at the center of the low- or high-dispersion photometric correction region. An extrapolation of the mean displacements is performed well outside the photometric correction limits to ensure that the bi-cubic spline fit done during the last stage of *CROSS-CORR* is well-behaved for all pertinent data regimes.

The actual displacements are then compared to the smoothed interpretation and any displacements that differ by 0.25 pixel from the smoothed array are modified to conform to the smoothed displacements. These final vector displacements are then expanded into a full 3-D ($768 \times 768 \times 2$) floating point displacement image (ΔSample and ΔLine) using a bi-cubic spline interpolation algorithm. The ΔSample values are located in the first level of the final image; the ΔLine values are located in the second level of the image.
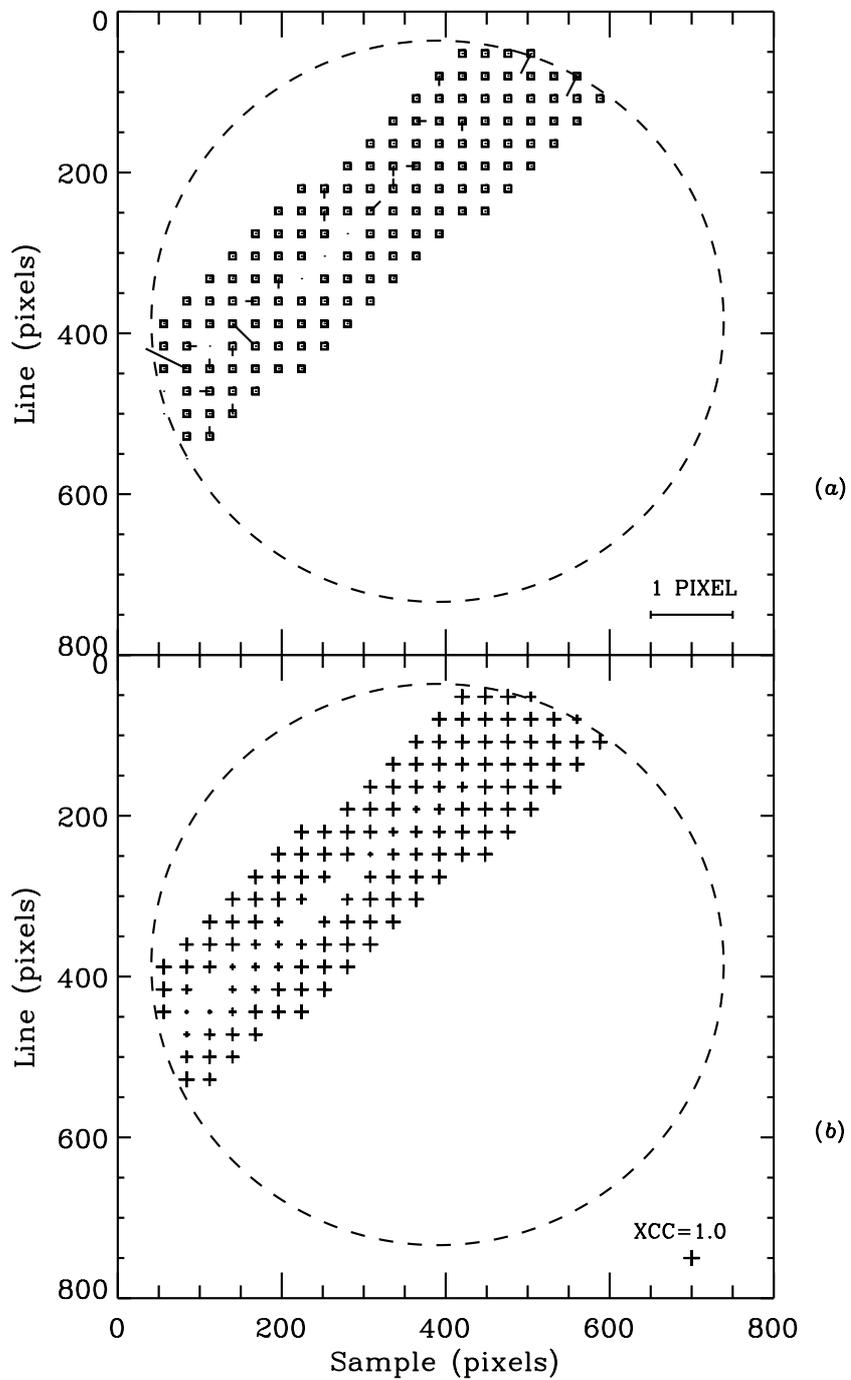
Figure 5.1: (*a*) Displacements computed between an SWP low-dispersion image and the SWP ITF. (*b*) Magnitude of the correlation coefficients corresponding to the results shown in Panel *a*.
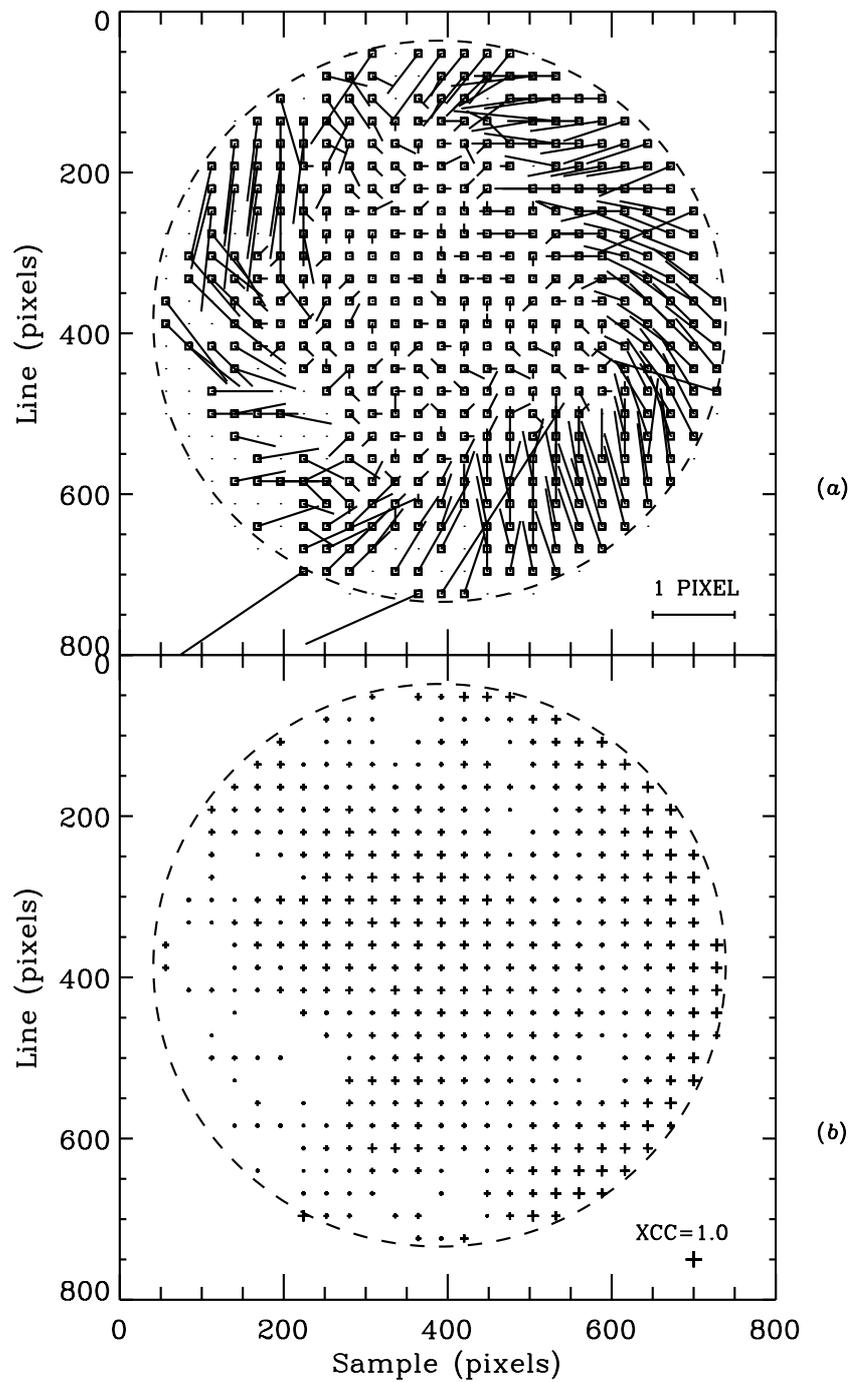
Figure 5.2: (*a*) Displacements computed between an SWP high-dispersion image and the SWP ITF. (*b*) Magnitude of the correlation coefficients corresponding to the results shown in Panel *a*.

5.5 *CROSS-CORR* **Output**

The VD data output by *CROSS-CORR* are an essential component for NEWSIPS processing. The information it carries is used as input to the photometric correction scheme to align properly the raw science image to the raw-space ITF image (Chapter 6) and by the geometric correction scheme as one of the vector components to the geometric resampling stage (Chapter 7). These specific data, however, are not retained directly in an *IUE* Final Archive output data product. As is discussed in Chapter 7, the VD FITS file (VDLO/VDHI) produced during the geometric correction procedure contains the summation of all the geometric corrections implemented to perform the single image resampling. The raw cross-correlation information can be retrieved indirectly from the binary table extension to the VDLO/VDHI. *CROSS-CORR* does not generate any ν flag values.

The *CROSS-CORR* module writes the following information to the HISTORY portion of the image label:

- window size,

- template size,

- ITF identification,

- percent successful correlations,

- median cross-correlation coefficient,

- standard deviation of the cross-correlation coefficient,

- mean computed displacement in pixels,

- maximum computed displacement in pixels, and

- number of isolated or semi-isolated displacements filtered.