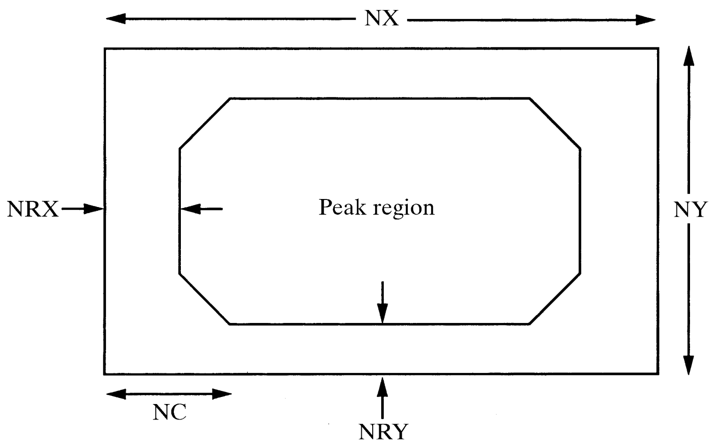

Fig. 11.2.4.1. The measurement-box definition used in *MOSFLM*. The measurement box has overall dimensions of NX by NY pixels (both odd integers). The separation between peak and background pixels is defined by the widths of the background rims (NRX and NRY) and the corner cutoff (NC). The size of the peak region is optimized separately for each of the standard profiles.

the background (background pixels) and those to be used for evaluating the intensity (peak pixels) are defined using a 'measurement box'. This is a rectangular box of pixels centred on the predicted spot position. Each pixel within the box is classified as being a background or a peak pixel (or neither). This mask can either be defined by the user, or the classification can be made automatically by the program. An example of a possible measurement-box definition is given in Fig. 11.2.4.1. The background parameters NRX, NRY and NC can be optimized automatically by maximizing the ratio of the intensity divided by its standard deviation, in a manner analogous to that described by Lehmann & Larsen (1974). It is generally assumed that the background can be adequately modelled as a plane, and the plane constants are determined using the background pixels. This allows the background to be estimated for the peak pixels, so that the background-corrected intensity can be calculated.

11.2.5. Integration by simple summation

11.2.5.1. *Determination of the best background plane*

The background plane constants a, b, c are determined by minimizing

$$R_1 = \sum_{i=1}^{n} w_i(\rho_i - ap_i - bq_i - c)^2, \qquad (11.2.5.1)$$

where ρ_i is the total counts at the pixel with coordinates $(p_i q_i)$ with respect to the centre of the measurement box, and the summation is over the n background pixels. w_i is a weight which should ideally be the inverse of the variance of ρ_i. Assuming that the variance is determined by counting statistics, this gives

$$w_i = 1/GE(\rho_i), \qquad (11.2.5.2)$$

where G is the gain of detector, which converts pixel counts to equivalent X-ray photons, and $E(\rho_i)$ is the expectation value of the background counts ρ_i. In practice, the variation in background across the measurement box is usually sufficiently small that all weights can be considered to be equal.

This gives the following equations for a, b and c, as given in Rossmann (1979),

$$\begin{pmatrix} \sum p^2 & \sum pq & \sum p \\ \sum pq & \sum q^2 & \sum q \\ \sum p & \sum q & n \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum p\rho \\ \sum q\rho \\ \sum \rho \end{pmatrix}, \qquad (11.2.5.3)$$

where all summations are over the n background pixels.

11.2.5.1.1. *Outlier rejection*

It is not unusual for the diffraction pattern to display features other than the Bragg diffraction spots from the crystal of interest. Possible causes are the presence of a satellite crystal or twin component, white-radiation streaks, cosmic rays or zingers. In order to minimize their effect on the determination of the background plane constants, the following outlier rejection algorithm is employed:

(1) Determine the background plane constants using a fraction (say 80%) of the background pixels, selecting those with the *lowest* pixel values.

(2) Evaluate the fit of all background pixels to this plane, rejecting those that deviate by more than three standard deviations.

(3) Re-determine the background plane using all accepted pixels.

(4) Re-evaluate the fit of all accepted pixels and reject outliers. If any new outliers are found, re-determine the plane constants.

The rationale for using a subset of the pixels with the lowest pixel values in step (1) is that the presence of zingers or cosmic rays, or a strongly diffracting satellite crystal, can distort the initial calculation of the background plane so much that it becomes difficult to identify the true outliers. Such features will normally only affect a small percentage of the background pixels and will invariably give higher than expected pixel counts. Selecting a subset with the lowest pixel values will facilitate identification of the true outliers. The initial bias in the resulting plane constant c due to this procedure will be corrected in step (3). Poisson statistics are used to evaluate the standard deviations used in outlier rejection, and the standard deviation used in step (2) is increased to allow for the choice of background pixels in step (1).

11.2.5.2. *Evaluating the integrated intensity and standard deviation*

The summation integration intensity I_s is given by

$$I_s = \sum_{i=1}^{m} (\rho_i - ap_i - bq_i - c), \qquad (11.2.5.4)$$

where the summation is over the m pixels in the peak region of the measurement box. If the peak region has mm symmetry, this simplifies to

$$I_s = \sum_{i=1}^{m} (\rho_i - c). \qquad (11.2.5.5)$$

To evaluate the standard deviation, this can be written as

$$I_s = \sum_{i=1}^{m} \rho_i - (m/n)\sum_{j=1}^{n} \rho_j, \qquad (11.2.5.6)$$

where the second summation is over the n background pixels.

The variance in I_s is

$$\sigma_{I_s}^2 = \sum_{i=1}^{m} \sigma_i^2 + (m/n)^2 \sum_{j=1}^{n} \sigma_j^2. \qquad (11.2.5.7)$$

From Poisson statistics this becomes