## 2.1. Target Segmentation and Clone Information Assignment

Given a typical microarray image with a regular pen-spotting pattern, many algorithms have been proposed to automatically detect the location of array within a scanned image and therefore determine the segmentation of the cDNA target location [Dougherty, 2001]. Some use a set of ″landing lights″ that form a predefined spotting pattern to assist the detection process. Assuming that a good robotic spotting process produces rigid grid pattern, we have chosen to implement a user-assisted graphic interface to overlay a grid for each image in order to assure that the grid is correctly placed. To satisfy the occasional need for grid refinement, one may activate the automatic refinement procedure before proceeding. The refinement procedure is performed as follows:

For each subarray,

a) Calculate the centers of detected cDNA spots, $(x_{s_i},\ y_{s_i})$ possessing high measurement quality;
b) Adjust the four corners of each grid by ±1, and then calculate the centers of segmentation $(x_{g_i},\ y_{g_i})$. Select an adjustment to minimized the following sum of differences:

$$\sum_i \sqrt{(x_{s_i} - x_{g_i})^2 + (y_{s_i} - y_{g_i})^2} \qquad (1)$$

c) Repeat step b until no adjustment is necessary.

Since the refinement procedure may be activated after the entire detection program has been executed, the statistics derived from the detection procedure, such as spot measurement quality, intensity or size, can be used to further qualify spots in the grid for the minimization process.

After segmentation, clone information (clone ID, titles, etc. are usually supplied with the location in a microtiter plate) must be connected to array locations before further processing. Information regarding non-printed, duplicated, negative-control, or housekeeping-gene locations is useful for normalization and other calibration purposes. The mapping of clone information to microarrays is determined by the robotic arrayer's deposition program. We use a standard array pattern that most arrayers can produce and a standard input format for other mass-fabricated microarray layouts. The standard array pattern is converted by the following procedure,

1. Let the pen configuration of the robot be $R_{pen}$×$C_{pen}$ (e.g, 2×2 or 2×4), and assume the robot always starts at A1, or the (1, 1) position from a $K$-well plate ($K$ = 96 or 384), and then moves first in the vertical direction (e.g., A1 to H1). The total number of robot pickups required to get to the $p$th plate at location $(r_{plate},\ c_{plate})$ is

$$n_p = \left\lfloor \frac{K\cdot(p-1)}{(R_{pen}\cdot C_{pen})} \right\rfloor + \left\lfloor \frac{(c_{plate}-1)}{C_{pen}} \right\rfloor \cdot \frac{R_{plate}}{R_{pen}} + \left\lfloor \frac{r_{plate}-1}{R_{pen}} \right\rfloor + 1 \qquad (2)$$

where $R_{plate}$ is the number of rows in the microtiter plate, and $\lfloor\bullet\rfloor$ denotes truncation to the integer. The first term in Eq. 2 is the number of pickups needed to exhaust one microtiter plate. Horizontal pickups and different starting positions can be similarly obtained.

2. Given the current robot pickup $n_p$, we can reverse map the corresponding plate location for a given pen ($r_{pent}$, $c_{pen}$) as

$$
\begin{aligned}
p &= \lfloor (n_p - 1)/(K / R_{pen} \cdot C_{pen}) \rfloor + 1 \\
c_{plate} &= \lfloor \alpha / (R_{plate} / R_{pen}) \rfloor + c_{pen} \\
r_{plate} &= \alpha \% (R_{plate} / R_{pen}) + r_{pen}
\end{aligned}
\qquad (3)
$$

where $\alpha = (n_p - 1)\%(K / R_{pen} \cdot C_{pen})$ and % denotes the modulus function.

3. Similarly, we can obtain a set of functions that maps from location at ($r_{array}$, $c_{array}$) to the deposition number, $n_d$, and vice versa, as follows:

$$n_d = (c_{array} - 1)R_{array} + r_{array} \qquad (4)$$

where each subarray consists of $R_{array}$ rows and $C_{array}$ columns. Also, we assume the deposition process starts at row 1 and column 1 (as typically defined as the up-left corner) and fills up the column first. Reversibly, given $n_d$ and a subarray position determined by a specific print-tip, ($r_{pen\_at}$, $c_{pen\_at}$), we can determine the coordinate of a particular deposition spot ($r_{array}$, $c_{array}$),

$$c_{array} = \lfloor (n_d - 1)/R_{array} \rfloor + 1, \quad r_{array} = (n_d - 1)\%R_{array} + 1 \qquad (5)$$

Upon obtaining Eqs. 2-5, we can convert plate position to and from array position by equating the number of depositions to the number of pickups, or $n_d = n_p$. For duplicated-spot printing, $n_d = 2n_p$.

4. Many other print-patterns are equivalent to Eqs. 2-5 by swapping row and column coordinates and/or changing the starting (1, 1) position.

The coordinates for each clone in an array are defined as ($r_{pen}$, $c_{pen}$, $r_{array}$, $c_{array}$). If the coordinates of a commercially available array are convertible to the stated coordinate system, then the mapping is a trivial task.

**References**

[Dougherty, 2001] *Microarrays: Optical Technologies and Informatics*, M. Bittner, Y. Chen and E. R. Dougherty, Edt. *Proceedings of SPIE*, Vol. 4266, 1-12, 2001.